Identification, heritability and relation with gene expression of novel DNA methylation loci for blood pressure

Identification, heritability and relation with gene expression of novel DNA methylation loci for blood pressure

Yisong Huang[1*], Miina Ollikainen[2, 3*], Maheswary Muniandy[2*], Tao Zhang[4], Jenny van Dongen[5], Guang Hao[1], Peter J. van der Most[6], Yue Pan[1], Natalia Pervjakova[7], Yan V. Sun[8], Qin Hui[8], Jari Lahti[9, 10], Eliza Fraszczyk[6], Xueling Lu[6, 30], Dianjianyi Su[11, 12], Melissa A. Richard[13], Gonneke Willemsen[5], Kauko Heikkila[2], Irene Mateo Leach[14], Nina Mononen[15], Mika Kähönen[16], Mikko A. Hurme[17], Olli T. Raitakari[18, 19], Amanda J Drake[20], Markus Perola[26], Marja-Liisa Nuotio[26], Yunfeng Huang[8], Batbayar Khulan[20], Katri Räikkönen[10], Bruce HR Wolffenbuttel[21], Alexandra Zhernakova[31], Jingyuan Fu[31, 32], Haidong Zhu[1], Yanbin Dong[1], Jana V. van Vliet-Ostaptchouk[6, 21, 27], Lude Franke[31], Johan G Eriksson[22, 23], Myriam Fornage[24, 25], Lili Milani[7], Terho Lehtimäki[15], Viola Vaccarino[8], Dorret I. Boomsma[5], Pim van der Harst[14, 33], Eco J. C. de Geus[5], Veikko Salomaa[26], Shengxu Li[28], Wei Chen[11], Shaoyong Su[1‡], James Wilson[29‡], Harold Snieder[6‡§], Jaakko Kaprio[2, 3‡§], Xiaoling Wang[1‡§]

Author Afflictions

[1] Georgia Prevention Institute, Medical College of Georgia, Augusta University, Augusta, GA, USA
[2] Institute for Molecular Medicine FIMM, HiLIFE, University of Helsinki, PO Box 20 (Tukholmankatu 8), Helsinki, Finland
[3] Department of Public Health, Faculty of Medicine, University of Helsinki, PO Box 20 (Tukholmankatu 8), Helsinki, Finland
[4] Department of Biostatistics, Shandong University School of Public Health, Jinan, China
[5] Department of Biological Psychology, Amsterdam Public Health research institute, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081BT, Amsterdam, The Netherlands
[6] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[7] Estonian Genome Center, Institute of Genomics, University of Tartu, 23 Riia Street, 51010, Tartu, Estonia
[8] Department of Epidemiology, Emory Rollins School of Public Health, Emory University, Atlanta, GA, USA
[9] Turku Institute for Advanced Studies, University of Turku, Turku, Finland
[10] Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland
[11] Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA
[12] Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, China
[13] Department of Pediatrics, Section of Hematology/Oncology, Baylor College of Medicine
[14] Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[15] Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland

[16] Department of Clinical Physiology, Finnish Cardiovascular Research Center – Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; Department of Clinical Physiology, Tampere University Hospital, Tampere 33521
[17] Department of Microbiology and Immunology, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland
[18] Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20520, Finland
[19] Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20014, Finland
[20] University/British Heart Foundation Centre for Cardiovascular Science, University of Edinburgh, Queen's Medical Research Institute, Edinburgh, UK
[21] Department of Endocrinology, University of Groningen, University Medical Center Groningen, the Netherlands.
[22] Department of General Practice and Primary health Care, Tukholmankatu 8 B, University of Helsinki, Finland
[23]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland
[24] Brown Foundation Institute of Molecular Medicine, Mc Govern Medical School, University of Texas Health Science Center at Houston
[25] Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston
[26] National Institute for Health and Welfare, P.O. Box 30, 00271 Helsinki, Finland
[27]Genomics Coordination Center, Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[28] Children's Minnesota Research Institute, Children's Hospitals and Clinics of Minnesota, Minneapolis, MN, USA
[29] Department of Physiology and Biophysics, University of Mississippi Medical Center, 2500 N. State St., Jackson, MS 39216 USA
[30] Laboratory of Environmental Medicine and Developmental Toxicology, Shantou University Medical College, 515041, Guangdong, China
[31] University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands
[32]Department of Pediatrics, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands
[33]Department of Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands

* These authors contributed equally to this work
‡ Senior authorships & these authors contributed equally to this work
§ Corresponding authors

Correspondence to
Harold Snieder, Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands, E-mail: h.snieder@umcg.nl

Jaakko Kaprio, Institute for Molecular Medicine FIMM, HiLIFE, University of Helsinki, PO Box 20 (Tukholmankatu 8), Helsinki, Finland, E-mail: jaakko.kaprio@helsinki.fi
Xiaoling Wang, Georgia Prevention Institute, Medical College of Georgia, Augusta, GA, USA, E-mail: xwang@augusta.edu

SHORT TITLE: BP related DNA methylation changes

Word count: 7145; **Tables:** 6

ABSTRACT

We conducted an epigenome-wide association study meta-analysis (meta-EWAS) on blood pressure (BP) in 4,820 individuals of European and African ancestry aged 14-69. Genome-wide DNA methylation data from peripheral leukocytes were obtained using the Infinium HumanMethylation450k BeadChip. The meta-EWAS identified 39 BP-related CpG sites with $p<1\times10^{-5}$. In silico replication in the CHARGE consortium of 17,010 individuals validated 16 of these CpG sites. Out of the 16 CpG sites, 13 showed novel association with BP. Conversely, out of the 126 CpG sites identified as being associated ($p<1x10^{-7}$) with BP in the CHARGE consortium, 21 were replicated in the current study. Methylation levels of all the 34 CpG sites that were cross-validated by the current study and the CHARGE consortium were heritable and 6 showed association with gene expression. Furthermore, 9 CpG sites also showed association with BP with $p<0.05$ and consistent direction of the effect in the meta-analysis of the Finnish Twin Cohort (199 twin pairs and 4 singletons; 61% monozygous) and the Netherlands Twin Register (266 twin pairs and 62 singletons; 84% monozygous). Bivariate quantitative genetic modeling of the twin data showed that a majority of the phenotypic correlations between methylation levels of these CpG sites and BP could be explained by shared unique environmental rather than genetic factors, with 100% of the correlations of systolic BP with cg19693031 (*TXNIP*) and cg00716257 (*JDP2*) determined by environmental effects acting on both systolic BP and methylation levels.

INTRODUCTION

Essential hypertension (EH) is a major health problem with global proportions. A report in the Lancet[1] estimated that in 2015 there were 1.13 billion people living with high blood pressure (BP) worldwide. While many pathways involved in the development of EH and corresponding treatment options have been discovered, the elevated blood pressure (BP) of 16 million hypertensive patients remains uncontrolled[2], indicating the need for further understanding of its pathogenesis. Epigenetics has recently been suggested as a massive regulatory machine that cannot be ignored in searching for the molecular understanding of EH [3, 4]. In fact, it may explain the late onset, progressive and quantitative nature of this disease better than variations in DNA sequence.

Epigenetic alterations of the genes of the renin-angiotensin-aldosterone system, a hormone system that is integral to the physiological regulation of BP, have been extensively tested in hypertensive animal models, providing one line of substantial evidence on the involvement of epigenetic regulation in the development of EH [5]. A recent, genome-wide, peripheral blood DNA methylation study in human by the CHARGE consortium[6], including a discovery and a replication panel, identified 13 CpG sites in or next to 8 genes that were differentially methylated in relation to BP. A methylation risk score based on these 13 CpG sites explained 1.4% and 2.0% of the inter-individual variation in systolic and diastolic BP, respectively. Expanding the methylation risk score to include 126 CpG sites that were Bonferroni significant ($p<1x10^{-7}$) in the overall meta-analysis did not explain additional phenotypic variance, indicating the need for further replication. Moreover, unlike sequence variation, epigenetic variation is influenced both by inherited and environmental factors[7, 8]. This is illustrated by the fact that 30-100% of the DNA methylation levels of the 13 BP-associated CpG sites described above is explained by heritable factors as estimated by the family data of the Framingham Heart Study [6]. However, the extent to which the link between BP and DNA methylation signatures is driven by inherited vs. environmental factors has not been investigated.

In the present meta-EWAS in leukocytes of 4,820 individuals of European (EA) and African ancestry (AA) aged 14-69, we first identified new DNA methylation signals associated with BP and validated these signals in the CHARGE consortium[6]; next we attempted to replicate the 126 previously identified signals by the CHARGE consortium [6] in our own meta-EWAS data; third we conducted twin modeling to estimate the heritability of DNA methylation correlated with BP, and finally we assessed the genetic and environmental sources of the correlation between DNA methylation and BP (Figure S1).

METHODS

Data availability

This study involves multiple cohorts. The genome wide DNA methylation data that support the findings of this study are available from the study PI of each cohort upon reasonable request and with permission of the Institutional Review Board of the universities where the participating cohort locates.

Study Populations

The discovery panel included 4820 individuals of EA and AA ancestries from 12 adult cohorts (average age ranges from 27.3 to 63.5 years old) and 2 youth cohorts (average age 16.2 and 17.7 years old) (Table 1). Details of each cohort are provided in the Supplemental Data. All studies obtained written informed consent from participants and were approved by local institutional review boards and ethics committees.

Blood Pressure Measurements

For all the cohorts, BP was measured after a period of rest and an average of 3 sequential readings was used as the phenotype for each analysis. For 3 cohorts (GSH, EpiGO and LACHY), BP was measured in a supine position, while for the other 11 cohorts, BP was measured in a sitting position. With the exception of the NTR cohort for which BP was measured within ±2 years from the methylation measurement, all the cohorts had BP measured concurrently with the collection of peripheral leukocytes for DNA methylation profiling. If antihypertensive medication was used, 15mmHg and 10mmHg were added to the measured SBP and DBP levels, respectively[9].

DNA Methylation Profiling

For all cohorts, genome-wide DNA methylation data were obtained from peripheral blood using the Illumina Infinium Human Methylation 450K Beadchip (Illumina Inc.). A detailed description on preprocessing and quality control steps for each cohort is provided in the Supplementary Data. For all cohorts, white blood cell sub-populations were estimated using the approach described by Houseman et al.[10].

Cohort Level Association Analysis

For cohorts only including unrelated subjects, a linear regression model was used to estimate the associations between DNA methylation (i.e. β values) and BP with methylation levels used as dependent variables adjusting for age, sex, ancestry (in samples including EAs and AAs), BMI, and white blood cell sub-populations. For cohorts including related subjects, a linear mixed effect model was used to account for sample relatedness.

Meta-Analyses & Cross Validation

Meta-analysis across the 14 cohorts was conducted using METAL[11] by converting the direction of effect and p-value observed in each cohort into a signed Z-score. CpG sites with $p \leq 1 \times 10^{-5}$ for either SBP or DBP were selected for replication in the CHARGE consortium (n=17,010). Replication was defined as consistent direction of the β-coefficient and FDR<0.05. Conversely, we also checked whether we could replicate the 126 CpG sites for BP identified in the overall meta-analysis of the CHARGE consortium in our own meta-EWAS results. Replication was again defined as a consistent direction of the β-coefficient and FDR<0.05.

Percent Variance Explained

Percent variance explained by the cross-validated BP associated CpG sites was calculated in the Lifelines DEEP cohort. To avoid overestimation of percent variance explained, this cohort was not included in the Meta-analysis. Percent variance explained by the cross-validated BP associated CpG sites is reported as the change in the adjusted R^2 from the model including these CpG sites compared to the model only including covariates (i.e. age, gender and BMI).

Pathway Analysis

Pathway enrichment analysis was conducted on the meta-analysis results of the genome-wide DNA methylation data using gene set enrichment analysis (GSEA)[12]. GSEA was performed on an unfiltered, ranked list of genes (ranked by the p values without consideration of directions), and a running-sum statistic was used to determine the enrichment of a priori defined gene sets (pathways) based on the gene ranks. All gene ontology biological process categories (c5.bp.v5.1) were assessed for enrichment at FDR<0.05. The CpG site showing the most significant p value within a gene was used to represent the DNA methylation level of the gene.

Associations of DNA Methylation and Gene Expression

Association tests of the cross-validated BP-associated CpGs (Figure S1) with transcripts that were located within 500kb distance of the corresponding CpGs were performed in the 391 individual twins of the Finnish Twin Cohort for whom both DNA methylation and gene expression data were available. Gene expression data were obtained using the Illumina Human HT-12 V4 expression Beadchip (Illumina, Inc, San Diego, CA) [13]. Linear mixed effects regression models were used with gene expression as the dependent variable, DNA methylation as the independent variable, age, sex, and BMI as fixed effects, and family as a random effect. An FDR < 0.05 was defined as significant association between DNA methylation and gene expression. BP associated gene expressions were defined as genes with their expression levels showing significant association with either SBP or DBP at $p < 0.05$.

Genetic and Environmental Determinants of DNA Methylation Associated with BP

For all the cross-validated BP associated CpG sites (Figure S1), we estimated the relative contributions of genetic and environmental factors to the variance of DNA methylation levels in the Finnish Twin Cohort and the Netherlands Twin Register using the R package OpenMx[14, 15]. Before analysis, age, sex and BMI were regressed out, and the DNA methylation residuals were used in the model fitting. Details of this univariate structural equation model for twin data (Figure S2) were described in a previous study [13]. In short, the model allows separation of the observed phenotypic variance into its genetic and environmental variance components including additive genetic variance (A), common environmental variance shared by a twin pair (C), and unique environmental variance specific to individuals (E). The monozygotic twins of each pair (MZ twins) have identical genome sequences, while dizygotic twins (DZ twins) share 50% of their segregating alleles. Shared environmental factors are exposures and experiences that affect co-twins similarly on average irrespective of zygosity, while unique environmental factors are the effects not shared by cotwins and include measurement error. Significance tests of individual variance components (A or C) were conducted by comparing full models with sub-models constraining paths from latent variables to trait values (a, c) to 0 using a chi-square test; as E contains measurement error, the significance of E is not tested. Statistical significance was defined as $p < 0.05$. The analysis was conducted in each twin cohort separately, then a meta-analysis across the two twin cohorts (FTC and NTR) was performed to estimate the mean heritability (h^2) using the Meta package in R [16]. Following the approach described by Asefa et al. [17], each h^2 was transformed using a logit function [18] and a random-effects model was used for the meta-analysis. The pooled h^2 (weighted by sample size) and 95% confidence intervals (CIs) were back-transformed. Heterogeneity between studies was quantified with Cochran's Q test and the I^2-statistic [19].

Sources Underlying the Associations between DNA Methylation and BP

For all the cross-validated BP associated CpG sites (Figure S1), we checked whether they were also significantly associated with BP in the meta-analysis of the Finnish Twin Cohort and the Netherlands Twin Register, i.e., a consistent direction of the β-coefficients and p<0.05. For those CpG sites significantly associated with BP in the meta-analysis of these two twin cohorts, we conducted bivariate structural equation modelling to test the extent to which the link between BP and DNA methylation was driven by genetic or environmental factors. Details of this model (Figure S3) have been described previously[13]. Briefly, the variation of DNA methylation and the variation of BP were decomposed into A, C, and E variance components. The bivariate model allows determination of the sources of the observed covariance between DNA methylation and BP by using a sequence of sub-models that test which genetic, shared environmental or unique environmental paths from DNA methylation to BP can be set to 0. For example, in Figure S3, if a_{21} (genetic path from DNA methylation to BP) cannot be set to 0, it means there is overlap between the genetic factors influencing DNA methylation and BP. The model further allows calculation of genetic and environmental correlations between the traits. Similar to the univariate analysis, the bivariate analysis was conducted in each twin cohort separately, then a meta-analysis was conducted to determine the genetic and environmental contributions to the correlation between DNA methylation and BP. Briefly, for each cohort, genetic (r_g) and environmental (r_e) correlations were calculated based on the variance/covariance matrix estimated from the bivariate twin modeling (Figure S3). The genetic contribution to the observed phenotypic correlation (r_{ph}) is a function of the heritability estimates of the two phenotypes and the r_g between them, i.e. $\sqrt{h_M^2} \times r_g \times \sqrt{h_{BP}^2}$. Similarly, the environmental contribution to r_{ph} is equal to $\sqrt{e_M^2} \times r_e \times \sqrt{e_{BP}^2}$. Then a random-effects model was used to estimate the meta-genetic and environmental contributions respectively with the 95% CIs.

RESULTS

The general characteristics of the study participants are listed in Table 1. A total of 4820 individuals were included from 14 cohorts with a wide range of mean SBP and DBP values. The prevalence of antihypertensive medication use also varied among the cohorts.

Our meta-analysis identified 39 CpG sites associated with SBP or DBP at $p<1\times10^{-5}$ (Manhattan and QQ plot, Supplementary Data Figure S4 and Figure S5; Table S1) with two CpG sites showing $p<1\times10^{-7}$. Out of these 39 CpG sites, the heterogeneity test across the cohorts reached significance (P<0.05) for five sites (cg06500161, cg00508575, cg19693031, cg12555233 & cg02711608, Supplementary Data Table S1). Further sensitivity tests by ancestry (EA vs AA) or age (adult cohorts vs. youth cohorts) did not support the heterogeneity being due to ancestry or age. Sixteen out of the 39 CpG sites including the three showing heterogeneity (cg06500161, cg00508575, & cg19693031) could be replicated (FDR<0.05) in the CHARGE consortium (Table 2). Of the 16 replicated CpG sites, only three (cg02711608, cg19693031, cg08857797) have previously been reported to be associated with BP (highlighted in gray in Table 2) and the other 13 were novel signals. Conversely, of the 126 CpG sites found to be associated with BP by the CHARGE consortium in the overall sample, 91 sites showed the same direction of effect in our meta-analysis with 21 sites having FDR < 0.05 (Table 3). These 21 CpG sites included the three CpG sites previously reported to be associated with BP (highlighted in gray in Table 3). In total, 34 CpG sites were cross-validated to be associated with BP by the current meta-analysis and the CHARGE consortium. To assess the impact of antihypertensive medication use, we stratified the meta-analysis of these 34

CpG sites by medication use and provide the results in Supplementary Data Table S2. In the individuals reporting no use of antihypertensive medications, the directions of the effects of all the CpG sites remained the same as for the overall sample with 27 out of the 34 retaining their significant associations ($p<0.05$) with BP, rendering it highly unlikely that the differentially methylated CpG sites we identified reflect drug treatment effects.

Inclusion of 33 out of the 34 CpG sites (cg02711608 was filtered out in the quality control step of the Lifelines DEEP cohort) explained an additional 3.31% and 3.99% of the interindividual variation in SBP and DBP, respectively, beyond the traditional BP covariates of age, gender and BMI in an additional sample from the Lifelines cohort (the Lifelines DEEP cohort, n=601) not included in the current meta-analysis. Details of this cohort are provided in the Supplemental Data. Using the Lifelines DEEP cohort, we further explored whether these 33 CpG sites were individually or collectively associated with EH. A total of 102 out of the 601 (16.97 %) participants were classified as having EH (i.e., SBP≥140 mmHg, or DBP≥90 mmHg, or taking antihypertensive medication). Out of the 33 CpG sites, 4 CpG sites (cg12593793, cg11376147, cg21766592 and cg06500161) were individually associated with EH with $p<0.05$ in the expected direction (Supplementary Data Table S3). Collectively, adding these 33 CpG sites in the model with age, gender and BMI as covariates increased Nagelkerke's pseudo R^2 from 34.3% to 47.7%.

Of the 34 CpG sites, the methylation levels of six sites were significantly associated with the expression of five genes in cis analysis (FDR <0.05) (Table 4). The methylation-gene expression associations did not differ by medication use. For all the CpG sites, increased methylation was associated with decreased gene expression (Supplementary Data Figure S6). Furthermore, expression of two genes (i.e. *ABCG1* and *LMNA*) showed significant association with BP. For both genes the direction of the association between CpG methylation and gene expression was as expected based on the association of CpG methylation and BP. For example, the methylation level of cg06500161 was negatively associated with *ABCG1* gene expression and positively associated with SBP. This was consistent with the negative association between *ABCG1* gene expression and SBP.

The pathway analyses yielded significant (FDR <0.05) enrichment of four biological process pathways for SBP-related DNA methylation changes, and six for DBP-related methylation changes in peripheral leukocytes (Supplementary Data Table S4). The primary pathway identified by the CHARGE consortium[6], the transport of neutral amino acids, also showed borderline significance in the current enrichment analyses for both SBP (FDR=0.060) and DBP (FDR=0.074).

The cohort-level results of the univariate structural equation model analysis on the DNA methylation levels of the 34 cross-validated CpG sites are listed in Table S5. For all the CpG sites, the best fitting models were AE models, with heritability estimates ranging from 31% to 83% in the Finnish Twin Cohort and 19% to 81% in the Netherlands Twin Register. The remaining part of the variation for DNA methylation of these CpG sites was attributable to environmental influences that are unique to the individual. Table 5 lists the heritability of the 34 CpG sites from the meta-analysis. The heritabilities ranged from 31% to 78%.

Of the 34 implicated CpG sites, the methylation level of 9 CpG sites showed significant association with BP in the meta-analysis of the Finnish Twin Cohort and the Netherlands Twin Register (Table 6). For these 9 sites, we estimated the relative contributions of genetic and environmental factors to the association between DNA methylation and BP. Given that the AE model has generally been the best

fitting model in previous twin studies of BP [20], which was again confirmed in the current study, the bivariate modeling was conducted using the AE model both for DNA methylation and BP. The cohort-level results are listed in Table S6 and the meta-analysis results are listed in Table 6. For the association of cg19693031 in the *TXNIP* gene and cg00716257 in the *JDP2* gene with SBP, the meta-analysis showed that the correlation due to the environmental contribution cannot be set to 0 while the genetic contribution can be set to 0, suggesting that the phenotypic correlation is determined by unique environmental factors in common to the two traits. Similar trends were observed for the association of cg11468085 with SBP and cg19693031 with DBP. For the associations of the other CpG sites with BP, both genetic and environmental contributions can be set to 0, indicating a larger sample size is needed to increase the power to distinguish the relative contributions of genetic and environmental factors to the observed phenotypic correlations.

DISCUSSION

In this epigenome-wide association study, we identified 13 novel CpG sites associated with BP and replicated 21 CpG sites previously identified in the overall meta-analysis of the CHARGE consortium[6]. We also showed that DNA methylation levels from 6 of the 34 cross-validated CpG sites were associated with gene expression. Although all of the 34 CpGs were heritable (31-78%), further bivariate twin modeling analyses in the Finnish Twin Cohort and the Netherlands Twin Register suggested that, among the 9 CpG sites that were associated with BP, the correlations of cg19693031(*TXNIP*) and cg00716257(*JDP2*) with SBP were primarily attributable to environmental factors that affect both traits, rather than genetic factors.

The 13 novel CpG sites that were associated with BP were annotated to 10 genes. Among these, only *ABCG1* and *ATP2B1* have previously been implicated in hypertension. For example, newly diagnosed hypertensive patients have been shown to have lower *ABCG1* expression in peripheral blood monocytes in comparison with normotensive controls[21]. This is consistent with our results in which we also observed that peripheral leukocyte *ABCG1* expression was negatively correlated with both SBP and DBP levels. Several genetic variants in *ATP2B1* have been associated with BP and hypertension in multiple GWA studies[22-24], and animal studies[25] have demonstrated that mice lacking *ATP2B1* in vascular smooth muscle cells had higher BP than wild type mice. In the current study, we observed higher methylation level of cg00508575 in *ATP2B1* associated with higher SBP level; however, the methylation status of this CpG site was not associated with *ATP2B1* expression levels in peripheral blood leukocytes. Further studies in other tissues such as vascular smooth muscle cells would be needed to clarify the functional role of this CpG site. The potential involvement of the other 8 genes in the pathogenesis of hypertension has not been directly addressed in the literature although some evidence is available on their involvement in cardiovascular diseases. For example, *TACC1* has been linked to inappropriate smooth muscle and endothelial cell proliferation in pulmonary arterial hypertension [26]. Several genome-wide association studies [27, 28] have reported variants in *DAB21P*, which encodes an inhibitor of cell growth and survival, that were associated with abdominal aortic aneurysm and atherosclerotic vascular diseases. *ALDH3B2*, encoding one member of the ALDH family of proteins that play a role in cell proliferation, differentiation, and responsiveness to environmental stress., has been suggested as a candidate gene for bisoprolol (a Beta blocker) responsiveness [29]. CpG sites in *MAN2A2* have been

associated with fasting insulin [30]. Further experimental validation of the role of these genes in BP regulation is warranted.

Similar to the CHARGE consortium[6] which involved cohorts from different ancestries (European, African American, and Hispanic) and a broad age range (18-80 years), the current study also included individuals from European and African American ancestry with an age range of 14-69 years. The fact that the signals could be cross-validated between these two studies, and that both studies showed the effect of the majority of BP related CpG sites to be homogeneous across the cohorts, ancestral groups, and different age groups, indicates that these BP-related CpG sites may be ethnicity- and age-independent. However, a clearer picture of the role of DNA methylation in the pathogenesis of EH in various age and population groups will require even larger EWASs spanning multiple age ranges and ancestry groups.

Interestingly, the majority of the 34 cross-validated CpG sites have been linked with other metabolic phenotypes including obesity, lipids, CRP, insulin resistance and type 2 diabetes by previous epigenome-wide association studies (Supplementary Data Table S7), indicating that DNA methylation may be one of the common factors related to the concurrence of multiple metabolic abnormalities. Indeed, epigenome-wide association studies have identified several CpG sites whose DNA methylation levels are associated with metabolic syndrome (MetS) including cg00574958 in the *CPT1A* gene[31] and cg06500161 in the *ABCG1* gene[32]. *ABCG1* cg06500161 has also been reported to be associated with fasting insulin [33], blood lipids[34], adiposity traits [35, 36] and type 2 diabetes [37, 38]. In the current study we observed for the first time that a higher methylation level of cg06500161 was also associated with higher BP levels. Taken together, these studies show that *ABCG1* cg06500161 is associated with each MetS component, though the causal direction of these associations has not been determined. Furthermore, although the other components of MetS can be viewed as consequences of obesity, the associations of these CpG sites with these MetS components are independent of obesity. Future studies are warranted with multivariate analyses targeting multiple metabolic traits to disentangle the mechanisms involved in the association of DNA methylation with MetS and its components.

Unlike genetic sequence variants, epigenetic variation is influenced by both genetic and environmental factors[7, 8]. We first quantified the genetic and environmental sources of the variation in the 34 cross-validated BP associated CpG sites and confirmed that the variance of all these 34 CpG sites was indeed determined by both genetics (31-78%) and environment (22-69%). Since BP is also a heritable trait, an interesting question is to what extent the link between BP and DNA methylation is driven by genetic or environmental factors in common to the two traits. We tried to answer this question using the Finnish Twin Cohort and the Netherlands Twin Register by conducting a bivariate twin modeling analysis on BP and the 9 CpG sites which showed association with BP in the meta-analysis of these two cohorts. Surprisingly, we observed that 100% of the correlations of BP with cg19693031(*TXNIP*) and cg00716257(*JDP2*) could be attributed to environmental factors in common to the two traits rather than genetic factors, despite evidence for high heritability of both methylation at those CpG sites and SBP. The apparent lack of shared genetic component indicates that the link between the methylation level of these two CpG sites and BP may be driven primarily by environmental conditions; the relatively modest sample size should be recognized however, and further confirmation is needed.

Our study has several limitations. First, it is cross-sectional, thus making it impossible to discern the temporal order between BP and DNA methylation. Second, the bivariate twin modeling analysis was only conducted in the Finnish Twin Cohort and the Netherlands Twin Register which included about 1000 twins. An even larger sample size is required to tease out reliably the relative contribution of genetic or environmental factors to the associations of BP with DNA methylation. Third, we did not conduct in vitro and in vivo functional studies to confirm the impact of these CpG sites on gene expression and subsequently on blood pressure, which is warranted for future research.

In conclusion, we identified 13 novel CpG sites associated with BP, and replicated several previously identified signals. These newly identified signals may aid in annotating the future gene findings by providing a potential molecular mechanism for BP regulation. Our study further provides new insights into the genetic and environmental sources of BP related DNA methylation signatures as well as their associations with BP. The identification of shared unique environmental factors rather than genetic factors between BP and DNA methylation of *TXNIP* and *JDP2* indicates that the environment plays a significant role in creating an association between DNA methylation signatures and BP.

SOURCES OF FUNDING

This study was supported by NIH/NHLBI (grant HL104125). Funding specific to each cohort was list below.

The Bogalusa Heart Study: The BHS is a joint effort of many investigators and staff members whose contribution is gratefully acknowledged. The current study was supported by grants R01AG016592 and R03AG060619 from National Institute of Aging, and P20GM109036 from the National Institute of General Medical Sciences of the National Institutes of Health.

The Georgia Stress and Heart Study: The GSH has been supported by NIH/NHLBI Grant HL69999. The current study was supported by NIH/NHLBI HL125577 & NIH/NHLBI HL104125. We would like to thank all the participants in the study and the staff at the Georgia Prevention Institute.

The DIetary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome Study: The DILGOM study was supported by the Finnish Academy (grant #118065). Dr. Salomaa was also supported by the Finnish Foundation for Cardiovascular Research.

The Emory Twin Study: The ETS has been supported by NIH R01HL68630, R01AG026255, R01HL125246, R01HL136205 and 2K24 HL077506. The current study was supported by NIH/NHLBI R21 HL106333-01A1.

The Estonian Genome Center of the University of Tartu: The EGCUT studies were financed by University of Tartu (grant "Center of Translational Genomics"), by Estonian Government (grant #SF0180142s08) and by European Commission through the European Regional Development Fund in the frame of grant "Centre of Excellence in Genomics" and Estonian Research Infrastructure's Roadmap and through FP7 grant #313010.

The Finnish Twin Cohort: The FTC has been supported by the Academy of Finland (grants 265240, 263278, 308248, 312073 to JK and 297908 to MO), and the Sigrid Juselius Foundation (to MO). The current DNA methylation study was supported by NIH/NHLBI grant HL104125.

The Helsinki Birth Cohort Study: The HBCS was supported by Finska Läkaresällskapet, the Finnish Special Governmental Subsidy for Health Sciences, Academy of Finland, Samfundet Folkhälsan, Liv och Hälsa, the Signe and Ane Gyllenberg Foundation, EU FP7 (DORIAN) project number 278603. A Drake and B Khulan were supported by a Scottish Senior Clinical Fellowship SCD/09 (to AD).

The Jackson Heart Study: The JHS is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I/HHSN26800001) and the University of Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung, and Blood Institute (NHLBI) and the National Institute for Minority Health and Health Disparities (NIMHD). The current DNA methylation study was supported by NIH/NHLBI grant HL104125. The authors also wish to thank the staff and participants of the JHS. Dr. Wilson is supported by U54GM115428 from the National Institute of General Medical Sciences.

The Lifelines Cohort Study: This work was supported by the National Consortium for Healthy Ageing (NCHA) (NCHA NGI Grant 050-060-810), a grant from the National Institute for Public Health and the Environment (RIVM) and the Ministry of Health, Welfare and Sports of the Netherlands (S/132005), the European Union's Seventh Framework programme (FP7/2007-2013) through the Biobank Standardisation and Harmonisation for Research Excellence in the European Union (BioSHaRE-EU) project, grant agreement 261433, and by the Dutch Diabetes Foundation (Diabetes Funds Junior Fellowship grant 2013.81.1673 to JVvVO).

The Netherlands Twin Register: Funding for the NTR was obtained from the Netherlands Organization for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW) grants 904-61-090, 985-10-002, 904-61-193,480-04-004, 400-05-717, Addiction-31160008, 016-115-035, Middelgroot-911-09-032, NWO-Groot 480-15-001/674, Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007 and 184.033.111); Spinozapremie (NWO- 56-464-14192), KNAW Academy Professor Award (PAH/6635) and University Research Fellow grant (URF) to DIB; Amsterdam Public Health research institute (former EMGO+) ; the European Science Foundation (ESF, EU/QLRT-2001-01254), the European Community's Seventh Framework Program (FP7- HEALTH-F4-2007-2013, grant 01413: ENGAGE and grant 602768: ACTION); the European Research Council (ERC Starting 284167, ERC Consolidator 771057, ERC Advanced 230374), Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the National Institutes of Health (NIH, R01D0042157-01A1, MH081802, DA018673, R01 DK092127-04, Grand Opportunity grants 1RC2 MH089951); the Avera Institute for Human Genetics, Sioux Falls, South Dakota (USA). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Computing was supported by NWO through grant 2018/EW/00408559, BiG Grid, the Dutch e-Science Grid and SURFSARA.

The Prevention of REnal and Vascular ENd stage Disease study: The PREVEND Study was funded by grants from the Dutch Kidney Foundation (grant E.033). The current DNA methylation study was supported by NIH/NHLBI grant HL104125.

The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular Research ; Finnish Cultural Foundation; The Sigrid Juselius Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; The Finnish Cultural Foundation - The Pirkanmaa Regional Fund; Diabetes Research Foundation of Finnish Diabetes Association; EU Horizon 2020 (grant 755320 for TAXINOMISIS); European Research Council (grant 742927 for MULTIEPIGEN project); Academy of Finland grant no. 322098; and Tampere University Hospital Supporting Foundation.

The EpiGenetic basis of Obesity induced cardiovascular disease and type 2 diabetes study: The EpiGO study was supported by NIH/NHLBI grant HL105689.

The Lifestyle, Adiposity, and Cardiovascular Health in Youth study: The LACHY study was supported by NIH/NHLBI grant HL64157. The current DNA methylation study was supported by NIH/NHLBI grant HL105689 & AHA grant GRNT20480211.

The Lifelines Deep Cohort: This project was funded by the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL) grant RP3 to LF

DISCLOSURE

VS has participated in a conference trip sponsored by Novo Nordisk and received an honorarium for participating in an advisory board meeting. He also has ongoing research collaboration with Bayer Ltd. (All unrelated to the present study)

REFERENCES

1. Collaboration NCDRF. Worldwide trends in blood pressure from 1975 to 2015: A pooled analysis of 1479 population-based measurement studies with 19.1 million participants. *Lancet*. 2017;389:37-55
2. Merai R, Siegel C, Rakotz M, Basch P, Wright J, Wong B, et al. Cdc grand rounds: A public health approach to detect and control hypertension. *MMWR. Morbidity and mortality weekly report*. 2016;65:1261-1264
3. Kato N, Loh M, Takeuchi F, Verweij N, Wang X, Zhang W, et al. Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. *Nature genetics*. 2015;47:1282-1293
4. Wang X, Falkner B, Zhu H, Shi H, Su S, Xu X, et al. A genome-wide methylation study on essential hypertension in young african american males. *PloS one*. 2013;8:e53938
5. Arif M, Sadayappan S, Becker RC, Martin LJ, Urbina EM. Epigenetic modification: A regulatory mechanism in essential hypertension. *Hypertension research : official journal of the Japanese Society of Hypertension*. 2019;42:1099-1113
6. Richard MA, Huan T, Ligthart S, Gondalia R, Jhun MA, Brody JA, et al. DNA methylation analysis identifies loci for blood pressure regulation. *American journal of human genetics*. 2017;101:888-902
7. Boks MP, Derks EM, Weisenberger DJ, Strengman E, Janson E, Sommer IE, et al. The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PloS one*. 2009;4:e6767
8. Gordon L, Joo JE, Powell JE, Ollikainen M, Novakovic B, Li X, et al. Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence. *Genome research*. 2012;22:1395-1406
9. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: Antihypertensive therapy and systolic blood pressure. *Stat Med*. 2005;24:2911-2935
10. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86
11. Willer CJ, Li Y, Abecasis GR. Metal: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26:2190-2191
12. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*. 2005;102:15545-15550
13. Huang Y, Ollikainen M, Sipila P, Mustelin L, Wang X, Su S, et al. Genetic and environmental effects on gene expression signatures of blood pressure: A transcriptome-wide twin study. *Hypertension*. 2018;71:457-464
14. Neale MC, Hunter MD, Pritikin JN, Zahery M, Brick TR, Kirkpatrick RM, et al. Openmx 2.0: Extended structural equation and statistical modeling. *Psychometrika*. 2016;81:535-549
15. Pritikin JN, Hunter MD, Boker S. Modular open-source software for item factor analysis. *Educational and psychological measurement*. 2015;75:458-474
16. Schwarzer G. Meta: An r package for meta-analysis. *R News*. 2007;7:40-45
17. Asefa NG, Neustaeter A, Jansonius NM, Snieder H. Heritability of glaucoma and glaucoma-related endophenotypes: Systematic review and meta-analysis. *Survey of ophthalmology*. 2019

18. Barendregt JJ, Doi SA, Lee YY, Norman RE, Vos T. Meta-analysis of prevalence. *Journal of epidemiology and community health*. 2013;67:974-978
19. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj*. 2003;327:557-560
20. Wang B, Liao C, Zhou B, Cao W, Lv J, Yu C, et al. Genetic contribution to the variance of blood pressure and heart rate: A systematic review and meta-regression of twin studies. *Twin research and human genetics : the official journal of the International Society for Twin Studies*. 2015;18:158-170
21. Xu M, Zhou H, Gu Q, Li C. The expression of atp-binding cassette transporters in hypertensive patients. *Hypertension research : official journal of the Japanese Society of Hypertension*. 2009;32:455-461
22. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, et al. Genome-wide association study of blood pressure and hypertension. *Nature genetics*. 2009;41:677-687
23. Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, et al. A large-scale genome-wide association study of asian populations uncovers genetic factors influencing eight quantitative traits. *Nature genetics*. 2009;41:527-534
24. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nature genetics*. 2018;50:1412-1425
25. Kobayashi Y, Hirawa N, Tabara Y, Muraoka H, Fujita M, Miyazaki N, et al. Mice lacking hypertension candidate gene atp2b1 in vascular smooth muscle cells show significant blood pressure elevation. *Hypertension*. 2012;59:854-860
26. Edgar AJ, Chacon MR, Bishop AE, Yacoub MH, Polak JM. Upregulated genes in sporadic, idiopathic pulmonary arterial hypertension. *Respiratory research*. 2006;7:1
27. Gretarsdottir S, Baas AF, Thorleifsson G, Holm H, den Heijer M, de Vries JP, et al. Genome-wide association study identifies a sequence variant within the dab2ip gene conferring susceptibility to abdominal aortic aneurysm. *Nature genetics*. 2010;42:692-697
28. Tragante V, Doevendans PA, Nathoe HM, van der Graaf Y, Spiering W, Algra A, et al. The impact of susceptibility loci for coronary artery disease on other vascular domains and recurrence risk. *European heart journal*. 2013;34:2896-2904
29. Hiltunen TP, Donner KM, Sarin AP, Saarela J, Ripatti S, Chapman AB, et al. Pharmacogenomics of hypertension: A genome-wide, placebo-controlled cross-over study, using four classes of antihypertensive drugs. *Journal of the American Heart Association*. 2015;4:e001521
30. Liu J, Carnero-Montoro E, van Dongen J, Lent S, Nedeljkovic I, Ligthart S, et al. An integrative cross-omics analysis of DNA methylation sites of glucose and insulin homeostasis. *Nature communications*. 2019;10:2581
31. Das M, Sha J, Hidalgo B, Aslibekyan S, Do AN, Zhi D, et al. Association of DNA methylation at cpt1a locus with metabolic syndrome in the genetics of lipid lowering drugs and diet network (goldn) study. *PloS one*. 2016;11:e0145789
32. Akinyemiju T, Do AN, Patki A, Aslibekyan S, Zhi D, Hidalgo B, et al. Epigenome-wide association study of metabolic syndrome in african-american adults. *Clinical epigenetics*. 2018;10:49
33. Hidalgo B, Irvin MR, Sha J, Zhi D, Aslibekyan S, Absher D, et al. Epigenome-wide association study of fasting measures of glucose, insulin, and homa-ir in the genetics of lipid lowering drugs and diet network study. *Diabetes*. 2014;63:801-807
34. Pfeiffer L, Wahl S, Pilling LC, Reischl E, Sandling JK, Kunze S, et al. DNA methylation of lipid-related genes affects blood lipid levels. *Circulation. Cardiovascular genetics*. 2015;8:334-342
35. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature*. 2017;541:81-86

36. Wang X, Pan Y, Zhu H, Hao G, Huang Y, Barnes V, et al. An epigenome-wide study of obesity in african american youth and young adults: Novel findings, replication in neutrophils, and relationship with gene expression. *Clinical epigenetics*. 2018;10:3
37. Chambers JC, Loh M, Lehne B, Drong A, Kriebel J, Motta V, et al. Epigenome-wide association of DNA methylation markers in peripheral blood from indian asians and europeans with incident type 2 diabetes: A nested case-control study. *The lancet. Diabetes & endocrinology*. 2015;3:526-534
38. Walaszczyk E, Luijten M, Spijkerman AMW, Bonder MJ, Lutgers HL, Snieder H, et al. DNA methylation markers associated with type 2 diabetes, fasting glucose and hba1c levels: A systematic review and replication in a case-control sample of the lifelines study. *Diabetologia*. 2018;61:354-368

Table 1. General characteristics of the study cohorts

Cohorts	N	Race [1]	Age, years Mean (SD)	Female, %	BMI, kg/m^2	SBP, mmHg	DBP, mmHg	HTN[4], %	AHT[5], %
Adult cohort									
BHS	968	EA, AA[2]	43.2(4.5)	56.5	30.8(7.5)	127.3(23.6)	81.2(14.6)	36.9	26.9
GSH	480	EA, AA[3]	27.3(3.5)	52.4	29.9(8.2)	114.9(13.3)	66.6(8.81)	9.39	4.80
DILGOM	512	EA	51.9(13.7)	53.7	26.9(4.8)	143.3(16.9)	83.4(10.0)	49.1	37.5
ETS	218	EA	55.7(3.4)	0	29.4(4.7)	135.66(18.4)	84.8(11.8)	56.3	34.4
EGCUT(Asthma)	173	EA	26.2(6.9)	64.2	22.8(3.0)	116.7(12.0)	73.2(9.4)	---[6]	---[6]
EGCUT (Young_Old)	100	EA	52.7(23.7)	52.0	26.7(5.1)	129.1(19.1)	79.6(10.4)	---[6]	---[6]
FTC	402	EA	62.3(3.7)	59.3	26.9(4.9)	150.2 (18.6)	87.8(11.0)	58.4	43.5
HBCS	159	EA	63.5(2.8)	0	27.5(3.8)	148.1(19.0)	91.0(10.3)	75.0	44.7
JHS	96	AA	38.4(4.3)	50	33.9(7.1)	127.8(23.7)	81.5(14.5)	50.0	34.3
Lifelines	150	EA	50.3(10.5)	58.7	28.0(5.1)	124.4(12.0)	74.1(8.60)	10.0	0.00
NTR	596[7]	EA	29.4(10.5)[8]	66.1	23.6(3.6)[8]	126.0(14.7)	76.3(10.5)	17.8	1.34
PREVEND	307	EA	46.7(10.0)	39.7	27(4.6)	131.1(20.8)	76.2(11.2)	48.5	19.5
YFS	188	EA	44.0(3.3)	38.8	26.3(4.4)	119.1(13.2)	73.2(9.5)	9.57	5.30
Youth cohort									
EpiGO	188	AA	17.7(1.7)	48.4	29.3(11.5)	114.8(15.1)	63.8(7.7)	9.57	0.00
LACHY	283	AA	16.2(1.3)	50.0	24.1(5.6)	112.9(10.1)	61.4(6.0)	3.53	0.00

BHS : the Bogalusa Heart Study; GSH: the Georgia Stress and Heart study; DILGOM: the DIetary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome Study; ETS: the Emory Twin Study; EGCUT: the Estonian Genome Center of the University of Tartu; FTC: the Finnish Twin Cohort; HBCS: the Helsinki Birth Cohort Study; JHS: the Jackson Heart Study; Lifelines: the Lifelines Cohort Study; NTR: the Netherlands Twin Register; PREVEND: Prevention of REnal and Vascular ENd stage Disease study; YFS: the Young Finns Study; EpiGO: the EpiGenetic basis of Obesity induced cardiovascular disease and type 2 diabetes study; and LACHY: the Lifestyle, Adiposity, and Cardiovascular Health in Youth study.

1) EA: European Ancestry; AA: African American.

2) For BHS, there are 70.3% EA and 29.7% AA.
3) For GSH, there are 52.5% EA and 47.5% AA.
4) HTN: hypertension; HTN definition for adults: SBP≥140mmHg or DBP≥90mmHg or on antihypertensive medication; HTN definition for youth: SBP≥95th or DBP≥95th percentile for age, sex and height.
5) AHT, antihypertensive treatment.
6) Only medication adjusted BP data were available.
7) This dataset included 499 MZ and 95 DZ twins as well as 2 spouses of twins. The 2 spouses of twins were excluded in the analyses only involving twins.
8) Age and BMI at blood sampling time.

Table 2. CpG sites showing association with BP in our analysis with p <1E-05 and replicated by CHARGE consortium with FDR<0.05

Probe ID	Chr.	Position	Gene	SBP				DBP			
				META		CHARGE		META		CHARGE	
				Direction*	p	Direction*	P	Direction*	p	Direction*	p
cg19693031	1	145441552	*TXNIP*	-	2.18E-07	-	3.10E-29	-	4.65E-05	-	1.80E-14
cg01343041	2	24397787	*C2orf84*	+	4.21E-07	+	2.30E-02	+	1.01E-04	+	5.74E-01
cg19695041	8	38615330	*TACC1*	-	6.26E-06	-	4.45E-05	-	3.36E-02	-	3.26E-03
cg13696706	9	124396830	*DAB2IP*	+	9.83E-08	+	3.95E-03	+	1.16E-03	+	7.90E-01
cg11468085	11	67435577	*ALDH3B2*	+	4.16E-06	+	1.75E-04	+	1.38E-04	+	2.27E-02
cg00508575	12	90050967	*ATP2B1*	+	6.47E-06	+	1.44E-03	+	9.20E-04	+	1.63E-01
cg05248321	14	20898128	*KLHL33*	+	7.01E-07	+	1.68E-05	+	1.88E-03	+	1.86E-03
cg02003183	14	103415882	*CDC42BPB*	+	3.66E-07	+	5.56E-07	+	1.54E-03	+	3.09E-02
cg12555233	15	91455366	*MAN2A2*	+	2.74E-06	+	6.25E-03	+	3.50E-03	+	2.97E-02
cg07558761	16	87866696	*SLC7A5*	+	8.83E-07	+	1.46E-05	+	1.71E-03	+	3.37E-04
cg07021906	16	87866833	*SLC7A5*	+	1.38E-06	+	1.65E-06	+	3.37E-03	+	5.73E-03
cg04583842	16	88103117	*BANP*	+	5.54E-09	+	4.16E-03	+	2.99E-06	+	8.28E-01
cg08857797	17	40927699	*VPS25*	+	9.64E-06	+	3.60E-10	+	4.98E-05	+	2.30E-06
cg02711608	19	47287964	*SLC1A5*	-	7.48E-06	-	2.00E-21	-	1.41E-03	-	4.30E-10
cg06500161	21	43656587	*ABCG1*	+	5.69E-06	+	1.01E-04	+	5.06E-05	+	1.01E-03
cg01820192	21	44869762	*C21orf125*	+	6.44E-06	+	1.66E-02	+	3.96E-02	+	4.75E-02

* "+" indicates that DNA methylation levels increase with BP increase. "-"indicates that DNA methylation levels decrease with BP increase.

CpG sites previously reported by the CHARGE consortium are highlighted in gray.

Table 3. Signals reported by CHARGE Consortium and replicated by the current study (FDR<0.05).

Probe ID	Chr.	Position	Gene	SBP				DBP			
				CHARGE		META		CHARGE		META	
				Direction*	p	Direction*	p	Direction*	p	Direction*	p
cg18933331	1	110186418	*Intergenic*	-	4.80E-09	-	7.64E-03	-	2.40E-08	-	1.17E-02
cg16246545	1	120255941	*PHGDH*	-	1.20E-22	-	4.11E-04	-	1.10E-09	-	4.34E-04
cg14476101	1	120255992	*PHGDH*	-	2.70E-34	-	4.27E-05	-	2.10E-21	-	2.31E-04
cg19693031	1	145441552	*TXNIP*	-	3.10E-29	-	2.18E-07	-	1.80E-14	-	4.65E-05
cg19266329	1	145456128	*Intergenic*	-	1.90E-12	-	3.61E-03	-	5.70E-05	-	2.22E-01
cg24955196	1	154982621	*ZBTB7B*	+	5.00E-08	+	8.28E-04	+	6.00E-06	+	5.09E-02
cg12593793	1	156074135	*Intergenic*	-	2.60E-12	-	3.22E-03	-	3.00E-07	-	7.49E-02
cg18119407	2	201980504	*CFLAR*	-	2.00E-09	-	7.10E-03	-	4.40E-05	-	4.51E-03
cg06690548	4	139162808	*SLC7A11*	-	1.60E-32	-	1.52E-05	-	7.90E-26	-	2.47E-05
cg18120259	6	43894639	*LOC100132354*	-	2.20E-21	-	5.58E-03	-	8.90E-14	-	4.55E-02
cg21429551	7	30635762	*GARS*	-	3.40E-16	-	6.84E-04	-	8.70E-06	-	1.34E-02
cg19390658	7	30636176	*GARS*	-	4.70E-09	-	2.58E-04	-	4.40E-06	-	3.30E-04
cg00008629	9	115093661	*ROD1*	-	6.50E-08	-	4.38E-03	-	8.00E-02	-	6.45E-02
cg11376147	11	57261198	*SLC43A1*	-	4.20E-21	-	6.66E-03	-	3.40E-12	-	1.34E-02
cg00574958	11	68607622	*CPT1A*	-	1.20E-13	-	1.04E-02	-	3.00E-10	-	6.83E-04
cg00716257	14	75897417	*JDP2*	-	6.00E-08	-	4.39E-03	-	4.40E-07	-	4.28E-01
cg26916780	15	64889554	*ZNF609*	-	4.50E-06	-	8.87E-03	-	3.70E-09	-	5.77E-02
cg08857797	17	40927699	*VPS25*	+	3.60E-10	+	9.64E-06	+	2.30E-06	+	4.98E-05
cg22304262	19	47287778	*SLC1A5*	-	1.40E-17	-	1.97E-05	-	9.60E-11	-	1.05E-02
cg02711608	19	47287964	*SLC1A5*	-	2.00E-21	-	7.48E-06	-	4.30E-10	-	1.41E-03
cg21766592	19	47288066	*SLC1A5*	-	2.60E-08	-	5.65E-04	-	1.10E-01	-	1.88E-02

* "+" indicates that DNA methylation levels increase with BP increase. "-"indicates that DNA methylation levels decrease with BP increase.

CpG sites that overlapped between Table 3 and Table 2 are highlighted in gray.

Table 4. Cross-validated CpG sites that show association with gene expression (±500kb) at FDR<0.05 in the Finnish Twin Cohort

DNAm ProbeID	DNAm annotation	GX ProbID	GX annotation	DNAm-GX			GX-SBP		GX-DBP		DNAm-SBP		DNAm-DBP	
				Dir.	P	FDR	Dir.	P	Dir.	P	Dir.	P	Dir.	P
cg14476101	*PHGDH*	240086	*PHGDH*	-	2.33E-11	1.19E-08	+	1.08E-01	+	5.62E-01	-	1.33E-02	-	2.88E-02
cg16246545	*PHGDH*	240086	*PHGDH*	-	2.94E-10	7.47E-08	+	1.08E-01	+	5.62E-01	-	5.23E-02	-	9.29E-02
cg06500161	*ABCG1*	6060377	*ABCG1*	-	1.37E-04	1.39E-02	-	2.75E-03	-	4.49E-04	+	4.74E-02	+	2.51E-01
cg12593793	*Intergenic*	6020424	*LMNA*	-	3.60E-04	2.28E-02	+	2.57E-04	+	1.78E-05	-	2.10E-01	-	2.53E-01
cg26916780	*ZNF609*	5960682	*RBPMS2*	-	8.61E-04	4.56E-02	+	1.88E-01	+	1.43E-01	-	6.84E-01	-	6.64E-01
cg02711608	*SLC1A5*	7610433	*SLC1A5*	-	8.98E-04	4.56E-02	+	4.62E-01	+	2.88E-01	-	4.34E-02	-	1.07E-01

GX: gene expression; Dir.: direction of the association

DNAm-GX: the association between DNA methylation and gene expression; GX-SBP: the association between gene expression and SBP; GX-DBP: the association between gene expression and DBP.

Table 5. Heritability of the 34 cross-validated CpG sites from the meta-analysis

Probeid	chr.	Position	Gene	h^2	95% CI
cg18933331	1	110186418	*intergenic*	0.72	0.62 - 0.81
cg16246545	1	120255941	*PHGDH*	0.77	0.69 - 0.84
cg14476101	1	120255992	*PHGDH*	0.74	0.63 - 0.82
cg19693031	1	145441552	*TXNIP*	0.56	0.53 - 0.60
cg19266329	1	145456128	*intergenic*	0.39	0.31 - 0.49
cg24955196	1	154982621	*ZBTB7B*	0.41	0.22 - 0.63
cg12593793	1	156074135	*intergenic*	0.65	0.37 - 0.86
cg01343041	2	24397787	*C2orf84*	0.65	0.62 - 0.68
cg18119407	2	201980504	*CFLAR*	0.36	0.33 - 0.39
cg06690548	4	139162808	*SLC7A11*	0.36	0.27 - 0.46
cg18120259	6	43894639	*LOC100132354*	0.63	0.56 - 0.69
cg21429551	7	30635762	*GARS*	0.68	0.65 - 0.71
cg19390658	7	30636176	*GARS*	0.35	0.32 - 0.38
cg19695041	8	38615330	*TACC1*	0.51	0.43 - 0.59
cg00008629	9	115093661	*ROD1*	0.78	0.65 - 0.87
cg13696706	9	124396830	*DAB2IP*	0.58	0.55 - 0.61
cg11376147	11	57261198	*SLC43A1*	0.41	0.25 - 0.59
cg11468085	11	67435577	*ALDH3B2*	0.54	0.51 - 0.57
cg00574958	11	68607622	*CPT1A*	0.44	0.30 - 0.58
cg00508575	12	90050967	*ATP2B1*	0.48	0.38 - 0.59
cg05248321	14	20898128	*KLHL33*	0.65	0.54 - 0.75
cg00716257	14	75897417	*JDP2*	0.31	0.11 - 0.62
cg02003183	14	103415882	*CDC42BPB*	0.62	0.59 - 0.65
cg26916780	15	64889554	*ZNF609*	0.40	0.34 - 0.46
cg12555233	15	91455366	*MAN2A2*	0.56	0.44 - 0.67
cg07558761	16	87866696	*SLC7A5*	0.53	0.38 - 0.68
cg07021906	16	87866833	*SLC7A5*	0.61	0.58 - 0.64
cg04583842	16	88103117	*BANP*	0.63	0.58 - 0.67
cg08857797	17	40927699	*VPS25*	0.48	0.21 - 0.75
cg22304262	19	47287778	*SLC1A5*	0.70	0.47 - 0.87
cg02711608	19	47287964	*SLC1A5*	0.69	0.61 - 0.75
cg21766592	19	47288066	*SLC1A5*	0.58	0.52 - 0.63
cg06500161	21	43656587	*ABCG1*	0.61	0.50 - 0.71
cg01820192	21	44869762	*C21orf125*	0.37	0.32 - 0.42

Table 6. Meta analysis results of the bivariate SEM analysis for BP and its associated CpG sites in FTC and NTR

Probeid	Trait	Gene	Association		Bivariate SEM analysis				
			Direction	P	r_{Ph}	G contribution	E contribution	$P_{G=0}$	$P_{E=0}$
cg19693031	SBP	*TXNIP*	--	0.0185	-0.047	0.023 (-0.040, 0.085)	**-0.070 (-0.132, -0.008)**	0.477	**0.026**
cg13696706	SBP	*DAB2IP*	++	0.0120	0.059	0.036 (-0.026, 0.098)	0.023 (-0.039, 0.085)	0.254	0.464
cg11468085	SBP	*ALDH3B2*	++	0.0148	0.026	-0.029 (-0.138, 0.079)	0.055 (-0.007, 0.116)	0.595	0.084
cg05248321	SBP	*KLHL33*	++	0.0117	0.056	0.034 (-0.028, 0.096)	0.022 (-0.04, 0.084)	0.279	0.492
cg00716257	SBP	*JDP2*	--	0.0317	-0.050	0.013 (-0.050, 0.075)	**-0.063 (-0.124, -0.001)**	0.692	**0.048**
cg04583842	SBP	*BANP*	++	0.0132	0.066	0.045 (-0.017, 0.107)	0.021 (-0.041, 0.083)	0.159	0.512
cg08857797	SBP	*VPS25*	++	0.0029	0.070	0.018 (-0.044, 0.08)	0.052 (-0.01, 0.114)	0.568	0.101
cg22304262	SBP	*SLC1A5*	--	0.0252	-0.050	-0.030 (-0.118, 0.058)	-0.020 (-0.082, 0.042)	0.509	0.535
cg06500161	SBP	*ABCG1*	++	0.0417	0.031	0.003(-0.098, 0.105)	0.028 (-0.034, 0.090)	0.952	0.369
cg19693031	DBP	*TXNIP*	--	0.0246	-0.062	-0.006 (-0.068, 0.056)	-0.056 (-0.118, 0.006)	0.857	0.076
cg11468085	DBP	*ALDH3B2*	++	0.0099	0.023	-0.013 (-0.126, 0.101)	0.036 (-0.037, 0.110)	0.828	0.333
cg08857797	DBP	*VPS25*	++	0.0022	0.090	0.065 (-0.016, 0.146)	0.025 (-0.037, 0.087)	0.117	0.432

r_{ph}: phenotypic correlation; G contribution: genetic contribution; E contribution: unique environmental contribution. A detailed explanation is provided in Supplementary Data Figure S2.

NOVELTY AND SIGNIFICANCE

What Is New?

- Identified 13 novel CpG sites of which their methylation levels are associated with BP
- Genetic factors contribute to the methylation variations of the BP associated CpG sites
- The phenotypic correlations between CpG sites and SBP are primarily attributable to environmental factors that affect both traits, rather than genetic factors.

What Is Relevant?

- The identification of shared unique environmental factors rather than genetic factors between BP and DNA methylation indicates that the environment plays a significant role in creating an association between DNA methylation signatures and BP.

Summary

In this study of 4,820 individuals of European and African ancestry aged 14-69, genome-wide DNA methylation data from peripheral leukocytes were obtained using the Infinium HumanMethylation450k BeadChip and blood pressures were measured during clinical visit. Linear regression or mixed models were used to identify differentially methylated CpG sites associated with BP. Univariate and bivariate structural equation modelings were used to further investigate to what extent the genetic and environmental factors influence DNA methylation and blood pressure in the Finnish Twin Cohort and the Netherlands Twin Register. Our study identify 13 more CpG sites with their methylation levels associated with BP and replicated 21 previously identified signals. Univariate twin modeling showed that genetic factors contributed to the methylation variations of all the 34 CpG sites with heritability estimates ranging from 31-78%. Bivariate twin modeling showed that 100% of the correlations of systolic BP with cg19693031 (*TXNIP*) and cg00716257 (*JDP2*) were determined by environmental effects acting on both systolic BP and methylation levels, rather than genetic factors.

Key words

Blood pressure, DNA methylation, Epigenome-wide association study, Twin study, Hypertension

Author Contributions

Original Study PI & Phenotyping: G. Willemsen, D. I. Boomsma, E. J. C. de Geus, B. HR. Wolffenbuttel, J. V. van Vliet-Ostaptchouk, J. Lahti, K. Raikkonen, J. G. Eriksson, T. Lehtimäki, M. A. Hurme, O. T. Raitakari, M. Kahonen, V. Salomaa, V. Vaccarino, K. Heikkilä, J. Kaprio

DNA methylation: I.M. Leach, P. van der Harst, L. Milani, A. Drake, B. Khulan, K. Raikkonen, N. Mononen, M. A. Hurme, H. Zhu, Y. Dong, M. Perola, M. Nuotio, A. Zhernakova, J. Fu, L. Franke

Data analysis: Y. Huang, M. Ollikainen, T. Zhang, M. Muniandy, J. van Dongen, G. Hao, P. J. van der Most, Y. Pan, Y. Sun, J. Lahti, E. Fraszczyk, Xu. Lu, D. Sun, N. Pervjakova, S. Su, M. A. Richard, B. Khulan, Q. Hui, Y. Huang.

Draft of the manuscripts: Y. Huang, M. Ollikainen, S. Su, J. Wilson, H. Snieder, J. Kaprio, X. Wang,

Primary PI for this study: L. Franke, M. Fornage, E. J. C. de Geus, J. V. van Vliet-Ostaptchouk, P. van der Harst, L. Milani, J. G. Eriksson, W. Chen, S. Li, T. Lehtimaki, S. Yan, S Su, J. Wilson, X. Wang, M. Ollikainen, J. Kaprio, H. Snieder

SUPPLEMENT

Identification, heritability and relation with gene expression of novel DNA methylation loci for blood pressure

Yisong Huang[1*], Miina Ollikainen[2, 3*], Maheswary Muniandy[2*], Tao Zhang[4], Jenny van Dongen[5], Guang Hao[1], Peter J. van der Most[6], Yue Pan[1], Natalia Pervjakova[7], Yan V. Sun[8], Qin Hui[8], Jari Lahti[9, 10], Eliza Fraszczyk[6], Xueling Lu[6, 30], Dianjianyi Su[11, 12], Melissa A. Richard[13], Gonneke Willemsen[5], Kauko Heikkila[2], Irene Mateo Leach[14], Nina Mononen[15], Mika Kähönen[16], Mikko A. Hurme[17], Olli T. Raitakari[18, 19], Amanda J Drake[20], Markus Perola[26], Marja-Liisa Nuotio[26], Yunfeng Huang[8], Batbayar Khulan[20], Katri Räikkönen[10], Bruce HR Wolffenbuttel[21], Alexandra Zhernakova[31], Jingyuan Fu[31, 32], Haidong Zhu[1], Yanbin Dong[1], Jana V. van Vliet-Ostaptchouk[6, 21, 27], Lude Franke[31], Johan G Eriksson[22, 23], Myriam Fornage[24, 25], Lili Milani[7], Terho Lehtimäki[15], Viola Vaccarino[8], Dorret I. Boomsma[5], Pim van der Harst[14, 33], Eco J. C. de Geus[5], Veikko Salomaa[26], Shengxu Li[28], Wei Chen[11], Shaoyong Su[1‡], James Wilson[29‡], Harold Snieder[6‡§], Jaakko Kaprio[2, 3‡§], Xiaoling Wang[1‡§]

Author Afflictions

[1] Georgia Prevention Institute, Medical College of Georgia, Augusta University, Augusta, GA, USA

[2] Institute for Molecular Medicine FIMM, HiLIFE, University of Helsinki, PO Box 20 (Tukholmankatu 8), Helsinki, Finland

[3] Department of Public Health, Faculty of Medicine, University of Helsinki, PO Box 20 (Tukholmankatu 8), Helsinki, Finland

[4] Department of Biostatistics, Shandong University School of Public Health, Jinan, China

[5] Department of Biological Psychology, Amsterdam Public Health research institute, Vrije Universiteit Amsterdam, Van der Boechorststraat 1, 1081BT, Amsterdam, The Netherlands

[6] Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

[7] Estonian Genome Center, Institute of Genomics, University of Tartu, 23 Riia Street, 51010, Tartu, Estonia

[8] Department of Epidemiology, Emory Rollins School of Public Health, Emory University, Atlanta, GA, USA

[9] Turku Institute for Advanced Studies, University of Turku, Turku, Finland

[10] Department of Psychology and Logopedics, University of Helsinki, Helsinki, Finland

[11] Department of Epidemiology, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, USA

[12] Department of Epidemiology and Biostatistics, School of Public Health, Peking University Health Science Center, Beijing, China

[13] Department of Pediatrics, Section of Hematology/Oncology, Baylor College of Medicine

[14] Department of Cardiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

[15] Department of Clinical Chemistry, Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33520, Finland

[16] Department of Clinical Physiology, Finnish Cardiovascular Research Center – Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland; Department of Clinical Physiology, Tampere University Hospital, Tampere 33521
[17] Department of Microbiology and Immunology, Faculty of Medicine and Health Technology, Tampere University, Tampere 33014, Finland
[18] Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku 20520, Finland
[19] Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku 20014, Finland
[20] University/British Heart Foundation Centre for Cardiovascular Science, University of Edinburgh, Queen's Medical Research Institute, Edinburgh, UK
[21] Department of Endocrinology, University of Groningen, University Medical Center Groningen, the Netherlands.
[22] Department of General Practice and Primary health Care, Tukholmankatu 8 B, University of Helsinki, Finland
[23]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland
[24] Brown Foundation Institute of Molecular Medicine, Mc Govern Medical School, University of Texas Health Science Center at Houston
[25] Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston
[26] National Institute for Health and Welfare, P.O. Box 30, 00271 Helsinki, Finland
[27]Genomics Coordination Center, Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
[28] Children's Minnesota Research Institute, Children's Hospitals and Clinics of Minnesota, Minneapolis, MN, USA
[29] Department of Physiology and Biophysics, University of Mississippi Medical Center, 2500 N. State St., Jackson, MS 39216 USA
[30] Laboratory of Environmental Medicine and Developmental Toxicology, Shantou University Medical College, 515041, Guangdong, China
[31] University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands
[32]Department of Pediatrics, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands
[33]Department of Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands

* These authors contributed equally to this work
‡ Senior authorships & these authors contributed equally to this work
§ Corresponding authors

Correspondence to
Harold Snieder, Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands, E-mail: h.snieder@umcg.nl

Jaakko Kaprio, Institute for Molecular Medicine FIMM, HiLIFE, University of Helsinki, PO Box 20 (Tukholmankatu 8), Helsinki, Finland, E-mail: jaakko.kaprio@helsinki.fi
Xiaoling Wang, Georgia Prevention Institute, Medical College of Georgia, Augusta, GA, USA, E-mail: xwang@augusta.edu

COHORT DESCRIPTION

Finnish Twin Cohort (FTC): In 2011, a comprehensive questionnaire was sent to all the living twins of the Finnish Twin Cohort born between 1945 and 1957 (11738 twins). A total of 8501 twins returned the questionnaire with a response rate of 72% [1]. Three questions in this questionnaire are related to hypertension: (1) When was your BP last measured? (2) Have you ever been told that you have elevated BP or hypertension? and (3) On how many days in the past year have you used antihypertensive medication? Based on the questionnaire, twins were defined as hypertensive if they had been diagnosed with hypertension or had taken antihypertensive medication for at least 2 months, while twins were defined as normotensive if they had not been diagnosed with hypertension and did not take antihypertensive medication. Based on the replies, 330 same-sex twin pairs free of self-reported previous diagnosis of myocardial infarction, congestive heart failure or stroke were defined as potentially discordant for hypertension. These twin pairs were phone interviewed and asked to participate in clinical assessments at the University of Helsinki during 2012-2015. A total of 222 twin pairs and 3 singletons (n=447) participated in the clinical protocol. Informed consent was obtained from each subject, and the study was approved by the Ethics Committee of the University Central Hospital of Helsinki.

During the clinical testing, the twins had a comprehensive physical examination, and their health history was recorded, including questions again about previous diagnoses of hypertension and use of anti-hypertensive medication. Resting BPs were measured four times during the visit, seated by a sphygmomanometer according to the JNC7 guidelines. The average of the last two readings of each measurement occasion was used to represent BP values. Information on medications was further complemented with data from community pharmacies. Based on the fact that only 50 twin pairs met the criteria for current discordance for hypertension (one twin on anti-hypertensive medication or with SBP≥140mmHg or DBP≥90mmHg and his/her co-twin not on anti-hypertensive medication, and with SBP<120mmHg and DBP<80mmHg), we included all the twins for the current analysis by using SBP and DBP as continuous variables. If antihypertensive medication was used, 15mmHg and 10mmHg were added to the measured SBP and DBP levels, respectively [2]. Fasting peripheral blood samples were obtained from 402 participants. Zygosity was determined by genotyping using Illumina HumanCoreExome BeadChip.

Georgia Stress and Heart study (GSH)[3, 4]: The GSH study is a longitudinal study, which was established in 1989 with the goal of exploring the development of cardiovascular risk factors. It consisted of 396 EA and 349 AA health youths aged from 7 to 16 years at baseline. All participants were recruited from the general population in the Augusta, Georgia area. The evaluation was conducted annually from 1989 to 2001 (visit 1-10), and every 1.5 years from 2002 to 2007 (visit 11-14). Most recently two more visits were conducted from 2008 to 2012 (visit 15-16). Blood samples were collected from visit 9. In the current study, all the participants in visit 15 with buffy coat available were selected for the current study (n=480).

Epigenetic Basis of Obesity Induced Cardiovascular Disease and Type 2 Diabetes (EpiGO) study [5, 6]. The EpiGO study was established in 2011 with the goal of identifying methylation changes involved in the pathogenesis of obesity and its related co-morbidities. The EpiGO study in total enrolled 351 obese (BMI percentile ≥95%) and 435 lean (BMI percentile ≤50%) youth aged 14 to 20 years with 70% of African-Americans (AA), 30% of European-Americans (EA), 47% of males and 53% of females. All the subjects were recruited from the Augusta, Georgia area from 2011-2015. The youths were free of any acute or chronic illness, and did not take daily prescription medication to treat diseases. Illumina 450k data were obtained from 192 African-American participants (96 obese vs. 96 lean, 50% of females). After quality control, genome-wide DNA methylation data for 188 subjects (96 obese vs. 92 controls) were included in the current analysis.

Lifestyle, Adiposity, and Cardiovascular Health in Youth (LACHY) study [5, 7]: The LACHY study was established in 2000 and finished in 2004 with the goal of determining the relations of fatness and fitness to cardiovascular disease (CVD) risk factors in the juvenile years. It consisted of 756 youths aged from 14 to 18 years recruited from general population, with roughly equal numbers of AA and EA of both genders. All participants were recruited from the Augusta, Georgia area. The youths were apparently healthy and did not take daily prescription medication to treat diseases. In this study, all AA participants with buffy coat available (n=286) were selected for the current study. After quality control, genome-wide DNA methylation data in leukocytes were available for 284 subjects.

Jackson Heart Study (JHS): The JHS is a large, community-based, longitudinal study whose participants were recruited from the Jackson metropolitan area. The primary objective of the JHS is to investigate the causes of EH and cardiovascular disease to learn how best to prevent and treat these diseases in AAs [8]. At baseline, a total of 5,302 AAs (21 to 94 years of age) were enrolled from 2000 to 2004. The participants have provided extensive medical and social history, and had an array of physical and biochemical measurements and diagnostic procedures. In particular, BP and buffy coat DNA were collected. For the current study, we selected 48 EH cases (50% females and 48% on antihypertensive medication) and 48 age (±2 years), sex and BMI (within the same category) matched normotensive controls for the genome-wide DNA methylation analysis. The including criteria for EH cases were: (1) age<45; (2) having genome-wide SNP data; and (3) on anti-hypertensive medication, or having SBP≥130mmHg or DBP≥85mmHg. The including criteria for normotensive controls were: (1) age<45; (2) having genome-wide SNP data; and (3) having SBP<120mmHg and DBP<80mmHg. To be in consistent with other cohorts, the current analysis used SBP and DBP as continuous variables rather than the phenotype of hypertension. If antihypertensive medication was used, 15mmHg and 10mmHg were added to the measured SBP and DBP levels, respectively [2].

Prevention of REnal and Vascular ENd stage Disease (PREVEND): The PREVEND is an ongoing large population-based longitudinal cohort started in 1997 and conducted in Groningen, the Netherlands. The main objective is to study microalbuminuria as a risk factor for renal and cardiovascular disease [9, 10]. During 1997-1998, all inhabitants of the city of Groningen between the ages of 28 and 75 yr (85,421 subjects) were asked to send in a morning urine sample and to fill out a short questionnaire. A total of 40,856 subjects (47.8%) responded. Subjects with a urinary albumin concentration of ≥10 mg/L (n = 7,768) together with a randomly selected control group with a urinary albumin concentration of <10 mg/L (n = 3,395) were invited for further investigations in an outpatient clinic. In total, 8,592 subjects completed the baseline test. Of the 8,592 subjects who completed the baseline test, 7928 Caucasians had buffy coat DNA available and BP measured. The current analysis included 2 genome-wide DNA methylation datasets from PREVEND. The first one included 98 EH cases (46.9% females and 35.7% on antihypertensive medication) and 98 age (±2 years), sex and BMI (within the same category) matched normotensive controls. The including criteria for EH cases were: (1) age<50; (2) having genome-wide SNP data; and (3) on anti-hypertensive medication, or having SBP≥140mmHg or DBP≥90mmHg. The including criteria for normotensive controls were: (1) age<50; (2) having genome-wide SNP data; and (3) having SBP<120mmHg and DBP<80mmHg. The second dataset consisted of 111 randomly selected samples, of which 28.3% were female and 20.9% used anti-hypertensive medication.

Emory Twin Study (ETS): The ETS consists of 283 middle-aged male Caucasian MZ and DZ twin pairs from the Vietnam Era Twin Registry who were born between 1946 and 1956[11]. All twins were examined in pairs at the Emory University General Clinical Research Center between 2002 and 2010. Genome wide DNA methylation data and measured BP were available in 218 twins aged 31-66[12]. In the current analysis we

included all the twins with epigenetic and BP data. If antihypertensive medication was used, 15mmHg and 10mmHg were added to the measured SBP and DBP levels, respectively [2].

Lifelines sample: Lifelines is a prospective population-based cohort to study health and health-related behaviours of 167,729 individuals living in the North of The Netherlands [13]. Details on clinical examination and biochemical measurements have been described elsewhere [13]. In short, a standardized protocol was used to obtain blood pressure and anthropometric measurements such as height, weight and waist circumference. Blood was collected in the fasting state, between 08:00 and 10:00 in the morning. Five groups were selected for genome-wide DNA methylation measurement from the baseline of the Lifelines study (all unrelated and European ancestry samples, n=13,436) [14] using the following criteria (n=50 for each group): (1) type 2 diabetes patients without CVD complications; (2) type 2 diabetes patients with CVD complications; (3) non-diabetic controls, with no history of CVD risk factors, and age- and sex- matched to groups 1 and 2; (4) healthy, obesity subjects (BMI≥30kg/m^2); and (5) healthy, normal-weight controls (BMI<25), and age- and sex-matched to group 4. The non-diabetic controls, healthy obesity subjects and normal-weight controls (n=150, i.e. group 3, 4, and 5) were included in the current analysis. [15]

Netherlands Twin Register (NTR) biobank project: For the NTR biobanking project, participants and collection procedures of blood, urine and buccal sample were described in detail previously[15]. The study also included parents of twins, siblings of twins and spouses of twins. In total, 3,264 blood samples from 3,221 NTR participants were assessed for genome-wide methylation, of which 3,089 samples from 3,057 subjects passed quality control. Blood pressure was measured in a subset of NTR participants as part of several projects that used similar methodology [16]. Here, we analyze SBP and DBP measured at rest. SBP and DBP were corrected for antihypertensive medication use by adding drug-class specific average treatment effects to the measured values [17, 18]. Only samples with good-quality DNA methylation data, white blood cell counts and BP measured ±2years from the methylation measurement were retained for EWAS and bivariate twin modeling analysis, leaving 596 subjects (age at the methylation measurement: 29.4±10.5; 66% females). This dataset included 499 MZ and 95 DZ twins as well as 2 spouses of twins. The univariate heritability analysis included 2802 twins (1736 MZ twins, and 1066 DZ twins) with good quality methylation data available. All subjects provided written informed consent. The study protocols were approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Centre, Amsterdam, an Institutional Review Board certified by the US Office of Human Research Protections (IRB number IRB-2991 under Federal-wide Assurance-3703; IRB/institute codes, NTR 03-180).

Dietary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome study (DILGOM): The Finnish study samples included a total of 513 unrelated individuals aged 25–74 years from the Helsinki area, recruited during 2007 as part of the DILGOM study, an extension of the FINRISK 2007 study described earlier[19]. Study participants were asked to fast overnight (at least 10 hours) prior to giving a blood sample. Genome-wide DNA methylation data were obtained from 512 individuals (Male= 237, Female=275, mean age 51.9 ± 13.7). The blood pressure measurements of DILGOM were carried out approximately three months earlier at baseline investigation of the FINRISK 2007 survey. As described[19], blood pressure measurements were performed by a trained nurse from the right arm of the participant who had been resting in a sitting position for five minutes before the first measurement. A standard mercury sphygmomanometer with a cuff size 14 x 36 cm was used. The first phase of the Korotkoff sounds was taken as the systolic pressure and the fifth phase as the diastolic pressure. Three measurements were performed and the mean of the last two measures was used in the analyses.

Estonian Genome Center of the University of Tartu (EGCUT) biobanking project: The EGCUT is the population-based database which comprises currently the health, genealogical and genome data of more than 51,530 individuals 18 years of age and up reflecting closely the age distribution in the adult Estonian

population. Participants of EGCUT recruited by the general practitioners (GP) from GP offices, physicians from the hospitals or data collectors from EGCUT's patient recruitment offices. Each participant filled out a Computer Assisted Personal interview during 1-2 hours at a doctor's or data collector's office, including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, quality of life). [20]Anthropometric and physiological measurements were also taken. Resting BPs were measured three times during the visit, seated by a sphygmomanometer according to the JNC7 guidelines. The average of three readings of each measurement occasion was used to represent BP values. If antihypertensive medication was used, 15mmHg and 10mmHg were added to the measured SBP and DBP levels, respectively. The current analysis included 2 genome-wide DNA methylation datasets from EGCUT. The first one included 91 asthma cases (mean age: 26.1±7.1 63.7% females) and 82 age- and gender-matched controls (mean age: 24.6±6.0, 64.6% females). Both cases and controls were required to have BMI<30kg/m^2 and to be non-smokers. The second one included 50 younger (age 22–34, 52% females) and 50 older (age 73–84, 52% females) individuals.

The Bogalusa Heart Study (BHS): The BHS, a long-term epidemiology study of the natural history of cardiovascular disease from childhood to adulthood, was founded by Dr. Gerald Berenson in 1973 [21]. The current study cohort consisted of 968 adult participants (680 whites and 288 blacks, 421 males, mean age=43.2±4.5 years) who had been recruited in the BHS from 2006 to 2010. Anthropometry and BP were measured in the morning. Participants' BP levels were obtained on right arms in a relaxed sitting position by 2 trained observers (3 times each). Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were recorded using a mercury sphygmomanometer. The mean values of the 6 readings were used for analysis. Whole blood samples were collected between 8:00 AM and 10:00 AM after overnight fasting for DNA methylation profiles assays. All participants in the BHS gave informed written consent. Study protocols were approved by the Institutional Review Board of Tulane University.

The Helsinki Birth Cohort (HBC): The HBCS comprises 13345 individuals (6370 women and 6975 men), born as singletons between 1934 and 1944 in one of the two main maternity hospitals in Helsinki and who were living in Finland in 1971 when a unique personal identification number was allocated to each member of the Finnish population. The HBCS, which has been described in detail elsewhere [22], has been approved by the Ethics Committee of the National Public Health Institute. Register data were linked with permission from the Finnish Ministry of Social Affairs and Health and the Finnish National Archives. In 2001–2004 at an average age of 61.5 years (s.d. = 2.9 and range = 56.7 –69.8 years), a randomly selected subsample of the cohort comprising 2003 individuals (1075 women and 928 men) was invited to a clinical examination including collection of a blood sample for (epi)genetic and biochemical studies. Genome-wide DNA methylation data were obtained from 83 evacuated men (mean age: 64±2.9 years) and 83 non-evacuated controls (mean age: 62.4±2.6 years) matched for sex, birth year and father's occupational status in childhood [23]. Blood pressure data were available for 164 participants. Blood pressure was measured from the right arm while the subject was in the sitting position, and it was recorded as the mean of 2 successive readings from a mercury sphygmomanometer. All measurements were performed by a team of 3 trained research nurses.

The Young Finns study (YFS): [24]YFS comprises a series of six cohorts, representing general population, born in 1962, 1965, 1968, 1971, 1974 and 1977 from five cities with university hospitals in Finland (Helsinki, Kuopio, Oulu, Tampere and Turku) [18]. A subsample of 188 (Male= 115, Female=73, mean age 44.0 ± 3.3) individuals were randomly assigned from a follow-up in 2011. The sample collection in 2011 is described in more detail elsewhere [34]. The categories of age in the methylation analysis were 40, 43, 46 and 49 years old, with group sizes of 50, 44, 55 and 35, in which 58 %, 68.2, 56.4 and 60 % were women, respectively. All of the participants were of western European descent. The study followed the guidelines of the Declaration

of Helsinki and was approved by the Ethical Review Committee of Turku University Hospital. All participants provided informed consent. Blood pressure from the right brachial artery was measured in the sitting position after a 5-minute rest with a random-zero sphygmomanometer (Hawksley & Sons Ltd, Lancin, United Kingdom) as described previously [25]. The average of three measurements was used in the analysis. Hypertension was defined as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, self-reported use of antihypertensive medication, or a self-reported hypertension diagnosis.

The Lifelines DEEP cohort: This cohort was not included in the current meta-analysis. It was used as an independent cohort to evaluate the contribution of the CpG sites to BP variation. A total of 601 adults (age: 45.6±13.1 years; 58.2% females, 9.48% taking antihypertensive medications) from the Lifelines DEEP cohort were included in this study based on available epigenome-wide methylation data and BP. If antihypertensive medication was used, 15mmHg and 10mmHg were added to the measured SBP and DBP levels, respectively. Initially, 1539 participants were enrolled in the Lifelines DEEP, which is a subpopulation of the Lifelines cohort in the north of The Netherlands [14, 26]. Except for the regular Lifelines procedures, additional deep molecular measurements were performed in Lifelines DEEP participants. All participants provided written informed consent. The Lifelines DEEP study was approved by the Medical Ethical Committee of the University Medical Center Groningen (UMCG), Groningen, The Netherlands.

GENOME-WIDE DNA METHYLATION PROFILING AND PROPROCESSING QUALITY CONTROL

The FTC, JHS, PREVEND, GSH, EpiGO, LACHY, YFS, DILGOM and ETS: The Minfi package [27] and CPACOR (incorporating Control Probe Adjustment and reduction of global CORrelation) package [28] were used for initial quantification, data preprocessing and quality control (QC). The key QC steps included: (1) Detectable probes were defined as the probes with detection P-value < $1x10^{-16}$ in more than 95% samples; (2) Detectable samples were defined as the samples with more than 95% CpG sites having a detection P-value <$1x10^{-16}$ and correct classification of gender based on the genome-wide DNA methylation data; (3) Probes on the X and Y chromosomes and the 65 SNP markers were excluded; (4) Illumina background correction and quantile normalization were applied to all intensity values and beta value was further calculated and used as the index of CpG methylation levels; (5) A principal component analysis (PCA) of the control probe intensities (excluding negative control probes) was performed and the resulting PCs up to 30 (different numbers were used by different cohorts) were stored; (6) White blood cell sub-populations were estimated using the approach described by Houseman et al. [29]; (7) A linear regression model was conducted for each CpG site with DNA methylation level as the dependent variable and the PCs from the control probe intensities as well as the estimated cell compositions as the independent variables. The residuals were calculated and used as indices of DNA methylation levels in further analysis. The above steps were conducted for each cohort separately.

BHS: The methylation data were normalized using the R package watermelon[30]. Given the importance of cellular heterogeneity in DNAm, we estimated the relative proportions of six pure cell types (CD4+ and CD8+ T-cells, natural killer cells, monocytes, granulocytes, and b-cells) in each sample using the R package Minfi package [27]. For correction of systematic technical biases in the 450K assay, the β scores were normalized using the dasen function, in which type I and type II intensities and methylated and unmethylated intensities are quantile-normalized separately after backgrounds equalization of type I and type II probes. Based on bead count and detection p-values, the following thresholds were set for removal: 1) samples having 1% of CpG sites with a detection p-value>0.05; 2) probes having 5% of samples with a detection p-value>0.05; 3) probes with bead count less than 3 in 5% of samples.

NTR: The methylation data were examined with the R-package MethylAid [31]. Only samples that passed all five quality criteria (using the default MethylAid thresholds) were kept for further analyses. Several probe-level QC steps were performed to filter out probes with low performance. For all samples, ambiguously mapped probes were excluded, based on the definition of an overlap of at least 47 bases per probe from Chen et al. [32], and all probes containing a SNP, identified in the Dutch population[33], within the CpG site (at the C or G position) were excluded, irrespective of minor allele frequency. For each sample individually, probes with an intensity value of zero (not present on the array of a particular sample), probes with a detection P value>0.01, and probes with a bead count <3 were excluded. After these steps, probes with a success rate <0.95 across samples were removed from all samples. The total number of CpGs after these filtering steps was 421,119. Only autosomal sites were kept in the current analyses (N=411,169).

EGCUT: Data pre-processing and quality control analyses were performed in R with the Bioconductor package minfi[27], using the original IDAT files extracted from the HiScanSQ scanner. 'Raw' pre-processing was used to convert the intensities from the red and the green channels into methylated and unmethylated signals. Beta values were computed using Illumina's formula [beta = M/(M + U + 100)]. The difference in the distribution of beta values for type I and type II probes was corrected using "SWAN", a normalization method to address systematic changes between type I and type II probes. Detection p-values were obtained for every CpG probe in each sample. Qualified CpG sites were defined as probes with detection P-value<0.01 in ≥95% of the samples. Qualified samples were defined as samples having ≥90% of CpG sites with detection P-value<0.01. Only autosomal sites were kept in the final analyses.

Lifelines: DNA methylation data were generated by the Genome Analysis Facility of UMCG, the Netherlands (www.rug.nl/research/genetics/genomeanalysisfacility/). Quality control (QC) and normalization steps are described in detail elsewhere [34] and in ESM Methods. In short, the QC pipeline by Touleimat and Tost was used with background correction and probe type normalization [35]. Then, normalized beta-values were log-transformed into M-values for downstream analysis, since they have been shown to perform better in studies with smaller sample sizes [36].

HBC: Quality control pipeline was set up using the R-package minfi, including intensity readouts, normalization, cell type composition estimation, β- and M-value calculation. We excluded samples with low intensity (badSampleCutoff<10.6) or deviant beta distribution based on visual inspection (n=5). We did not detect any gender discrepancy. Data were normalized with functional normalization (FunNorm). Of the probes, we excluded those with detection p-value >0.01 in >50% of samples (ref. Minfi), non-autosomal and non-specific binding probes as well as probes with SNPs in the interval for which the Illumina probe is designed to hybridize and if they were located close (10bp from query site) to a SNP which had a minor allele frequency of ≥0.05. Probes located in the X and Y chromosome were also excluded. This yielded a total number of probes of 424,844. Batch effects were identified by inspecting the association of principal components of the methylation levels with possible technical batches using linear regressions and visual inspection of PCA plots using the Bioconductor R package *shinyMethyl* (version 0.99.3). Identified batch effects (i.e. array column) were removed using the Empirical Bayes' (EB) method *ComBat* [37]. Batch corrected beta values after *ComBat* were used for all further statistical analyses and white blood cell sub-populations were estimated using the Houseman et al [29].

The Lifelines Deep Cohort: For genome-wide DNA methylation data, 500 ng of genomic DNA was bisulfite-converted using the EZ DNA Methylation kit (Zymo Research Corp., USA) and hybridized on Illumina Human Methylation 450K BeadChip arrays (Illumina, Inc.) according to the manufacturer's protocols. The original IDAT files were generated by the Illumina iScan BeadChip scanner and performed by the Human Genotyping facility (HugeF) of ErasmusMC, The Netherlands (http://www.glimDNA.org/). Quality control and normalization details are described elsewhere (unpublished manuscript). Briefly, the R-package "minfi"

was used to check on the quality of probes and samples. By following the Touleimat and Tost's pipeline [35], we performed background correction and probe type normalization using preprocessQuantile implemented in the minfi package [27]. Methylation levels at each CpG site were expressed as the ratios of the methylated intensity over the total intensity (β-values). Normalized M-values were used for all subsequent statistical analyses, while β-values were used for biological interpretation in tables and figures.

FIGURE LEGEND

Figure S1. Flow diagram of the study design.

Figure S2. Path diagram for univariate structural equation model

The observed phenotypes (P) for co-twins are shown in squares, and latent factors are shown in circles. Correlation between genetic factors (A) is 1 in monozygotic (MZ) twin pairs and 0.5 in dizygotic (DZ) twin pairs. Correlation between common (or shared) environmental factors (C) is 1 for MZ/DZ twin pairs. Unique environmental factors (E) are always uncorrelated.

Figure S3. Path diagram for bivariate structural equation model (Cholesky model)

The observed two phenotypes (P1, P2) for co-twins are shown in squares, and latent factors are shown in circles. Similar to the univariate model, correlation between genetic factors of the same phenotype (A1, A2) is 1 in monozygotic (MZ) twin pairs and 0.5 in dizygotic (DZ) twin pairs, while correlation between common (or shared) environmental factors (C) is 1 for both MZ/DZ twin pairs. Unique environmental factors (E) are always uncorrelated between two twins. The paths of genetic factors on each phenotype are a_{11} and a_{22} respectively, and the path of genetic factors between two phenotypes is a_{21}. Similarly, The paths of common and unique environmental factors on each phenotype are c_{11} and c_{22}, and e_{11} and e_{22}, respectively, and the path of common and unique environmental factors between two phenotypes are c_{21} and e_{21}. As the correlation is the standardized covariance, we can calculate the genetic correlation (r_g) based on the variance/covariance matrix estimated from the bivariate twin modeling, i.e. $r_g = \frac{(a_{11} \times a_{21})}{\sqrt{a_{11}^2 \times (a_{21}^2 + a_{22}^2)}}$, and the common and unique environmental correlations (r_c and r_e) as well, i.e. $r_c = \frac{(c_{11} \times c_{21})}{\sqrt{c_{11}^2 \times (c_{21}^2 + c_{22}^2)}}$ and $r_e = \frac{(e_{11} \times e_{21})}{\sqrt{e_{11}^2 \times (e_{21}^2 + e_{22}^2)}}$. The genetic contribution to the observed phenotypic correlation (r_{ph}) is a function of both heritabilities of the two phenotypes and the r_g between them, i.e. $\sqrt{h_1^2} \times r_g \times \sqrt{h_2^2}$. Similarly, the environmental contributions to the observed correlation are equal to $\sqrt{c_1^2} \times r_c \times \sqrt{c_2^2}$ and $\sqrt{e_1^2} \times r_e \times \sqrt{e_2^2}$. Dividing these terms by r_{ph} yields the proportion of r_{ph} due to genetic, common and unique environmental factors, repectively, because $r_{ph} = \sqrt{h_1^2} \times r_g \times \sqrt{h_2^2} + \sqrt{c_1^2} \times r_c \times \sqrt{c_2^2} + \sqrt{e_1^2} \times r_e \times \sqrt{e_2^2}$.

Figure S4. Manhattan plot of the genome-wide DNA methylation analysis. The blue line indicates the threshold of P<1E-05. Figure S4a: SBP; Figure S4b: DBP.

Figure S5. QQ plot of the genome-wide DNA methylation analysis. Figure S5a: SBP; Figure S5b: DBP.

Figure S6. Scatter plot of the assocations between DNA methylation and gene expression. X axis is DNA methylation level, which are the residuals of beta values after regressing out the effect of cell composition and 5 PCs from the control probes. Y axis is gene expression level, which are the residuals of the gene expression levels after regressing out the batch effect.

Table S1. CpG sites associated with SBP or DBP at $P<1\times10^{-5}$ from our meta-analysis.

Probe ID	Chr.	Gene	SBP			DBP		
			Direction	P value	Het P value	Direction	P value	Het P value
cg06500161	21	*ABCG1*	+	5.69E-06	5.79E-03	+	5.06E-05	3.45E-02
cg11468085	11	*ALDH3B2*	+	4.16E-06	2.32E-01	+	1.38E-04	2.22E-01
cg00508575	12	*ATP2B1*	+	6.47E-06	4.50E-02	+	9.20E-04	2.62E-01
cg04583842	16	*BANP*	+	5.54E-09	2.91E-01	+	2.99E-06	2.33E-01
cg05166473	16	*BANP*	+	1.26E-06	4.31E-01	+	8.71E-04	4.45E-01
cg01820192	21	*C21orf125*	+	6.44E-06	9.82E-01	+	3.96E-02	5.67E-01
cg14562076	21	*C21orf125*	+	2.93E-06	4.35E-01	+	3.38E-03	2.41E-01
cg01343041	2	*C2orf84*	+	4.21E-07	5.77E-01	+	1.01E-04	1.08E-01
cg18587476	2	*C2orf84*	+	3.98E-06	6.66E-01	+	3.59E-03	2.51E-01
cg19849557	6	*C6orf136*	+	2.64E-06	7.81E-01	+	4.07E-02	5.50E-01
cg02003183	14	*CDC42BPB*	+	3.66E-07	2.85E-01	+	1.54E-03	9.01E-02
cg10258505	14	*CDC42BPB*	+	3.95E-06	8.64E-01	+	2.43E-04	2.80E-01
cg23570433	11	*CTR9*	+	2.52E-03	6.34E-01	+	3.07E-06	3.14E-01
cg13696706	9	*DAB2IP*	+	9.83E-08	5.55E-01	+	1.16E-03	8.16E-01
cg10785537	2	*DCTN1*	+	6.50E-06	3.81E-01	+	2.43E-02	3.90E-01
cg24461627	10	*DIP2C*	+	2.35E-05	9.57E-01	+	6.49E-06	3.78E-01
cg19637821	12	*FBRSL1*	+	4.98E-07	6.38E-01	+	2.15E-02	2.08E-01
cg22590032	5	*FLT4*	+	3.94E-06	2.03E-01	+	3.06E-03	6.15E-01
cg10342963	15	*IGF1R*	+	2.91E-07	4.26E-01	+	6.48E-05	6.32E-01
cg05248321	14	*KLHL33*	+	7.01E-07	8.70E-01	+	1.88E-03	9.76E-01
cg08059112	19	*LINGO3*	+	2.42E-06	6.52E-01	+	1.85E-03	6.28E-01
cg12555233	15	*MAN2A2*	+	2.74E-06	4.87E-02	+	3.50E-03	2.90E-01
cg03835709	12	*NAV3*	+	8.75E-06	1.21E-01	+	8.24E-03	6.20E-01
cg13673536	9	*PIP5K1B*	+	1.73E-04	3.32E-01	+	4.24E-07	8.25E-01
cg20738719	3	*SEMA5B*	+	2.23E-07	3.36E-01	+	2.23E-04	7.02E-01
cg26401492	3	*SFMBT1*	+	1.08E-02	1.15E-02	+	5.92E-06	5.16E-01
cg02711608	19	*SLC1A5*	-	7.48E-06	2.50E-03	-	1.41E-03	3.36E-02
cg07021906	16	*SLC7A5*	+	1.38E-06	6.98E-01	+	3.37E-03	4.38E-01
cg07558761	16	*SLC7A5*	+	8.83E-07	1.90E-01	+	1.71E-03	8.45E-02
cg04690793	6	*SNRPC*	+	3.98E-06	9.09E-01	+	1.57E-03	4.90E-01
cg02980023	3	*STAB1*	+	3.78E-07	2.84E-01	+	5.86E-06	5.31E-01
cg19695041	8	*TACC1*	-	6.26E-06	8.09E-01	-	3.36E-02	5.34E-01
cg19693031	1	*TXNIP*	-	2.18E-07	1.23E-03	-	4.65E-05	7.63E-02
cg08857797	17	*VPS25*	+	9.64E-06	6.44E-01	+	4.98E-05	7.75E-01
cg00916854	12	*VPS37B*	+	7.99E-06	8.48E-02	+	2.06E-03	1.20E-01
cg09012734	21	*WDR4*	+	3.38E-02	2.52E-01	+	8.42E-06	2.21E-01
cg00989229	16	*Intergenic*	+	1.37E-02	6.58E-01	+	8.30E-06	3.47E-01
cg03411579	12	*Intergenic*	+	9.35E-07	2.54E-01	+	2.68E-04	2.19E-01
cg08232160	16	*Intergenic*	+	9.11E-06	1.86E-01	+	5.15E-03	5.82E-03

Table S2. Stratified analysis of the 34 CpG sites in participants with or without antihypertensive medication

			Without medication (n=3855)				With medication (n=790)			
			SBP		DBP		SBP		DBP	
Probe ID	Chr.	Gene	Direction	P value	Direction	P value	Direction	P value	Direction	P value
cg18933331	1	*intergenic*	-	**2.16E-02**	-	**1.89E-02**	+	6.17E-01	+	9.88E-01
cg16246545	1	*PHGDH*	-	**5.20E-04**	-	**1.14E-02**	-	5.78E-01	-	7.89E-01
cg14476101	1	*PHGDH*	-	**1.20E-04**	-	**7.11E-03**	-	7.23E-01	-	6.20E-01
cg19693031	1	*TXNIP*	-	**1.23E-02**	-	**4.82E-04**	-	1.20E-03	+	7.17E-01
cg19266329	1	*intergenic*	-	1.93E-01	-	9.85E-01	-	4.01E-01	+	7.11E-02
cg24955196	1	*ZBTB7B*	+	**1.54E-02**	+	**4.18E-03**	+	5.80E-03	+	3.00E-02
cg12593793	1	*intergenic*	-	1.66E-01	-	3.26E-01	-	5.52E-03	-	5.47E-01
cg01343041	2	*C2orf84*	+	**2.18E-02**	+	9.93E-02	+	9.19E-03	+	1.65E-01
cg18119407	2	*CFLAR*	-	**4.32E-02**	-	**1.38E-03**	-	3.72E-01	+	3.60E-01
cg06690548	4	*SLC7A11*	-	**2.21E-03**	-	**1.61E-04**	-	5.32E-01	-	8.79E-02
cg18120259	6	*LOC100132354*	-	7.20E-02	-	6.90E-01	-	7.03E-01	-	5.75E-01
cg21429551	7	*GARS*	-	**1.20E-02**	-	**3.43E-02**	-	9.55E-01	+	8.10E-01
cg19390658	7	*GARS*	-	**8.64E-04**	-	**2.48E-02**	-	9.86E-01	+	7.45E-01
cg19695041	8	*TACC1*	-	**2.97E-04**	-	5.92E-01	-	5.55E-02	-	6.36E-03
cg00008629	9	*ROD1*	-	2.27E-01	-	9.84E-01	-	1.85E-02	-	1.80E-01
cg13696706	9	*DAB2IP*	+	**9.16E-06**	+	**1.39E-02**	+	3.52E-04	+	8.21E-02
cg11376147	11	*SLC43A1*	-	**1.88E-02**	-	**1.83E-02**	-	8.51E-01	+	9.60E-01
cg11468085	11	*ALDH3B2*	+	**9.01E-05**	+	**2.04E-03**	+	2.69E-01	+	4.39E-01
cg00574958	11	*CPT1A*	-	4.65E-01	-	9.28E-02	-	9.61E-01	-	4.11E-01
cg00508575	12	*ATP2B1*	+	**5.45E-03**	+	**3.01E-02**	+	8.62E-04	+	1.43E-01
cg05248321	14	*KLHL33*	+	**2.87E-04**	+	**6.44E-03**	+	4.41E-02	+	3.78E-01
cg00716257	14	*JDP2*	-	6.42E-02	-	6.81E-01	-	7.37E-01	-	2.68E-01
cg02003183	14	*CDC42BPB*	+	**1.44E-05**	+	5.92E-02	+	4.18E-01	-	2.39E-01
cg26916780	15	*ZNF609*	-	1.56E-01	-	2.27E-01	-	5.27E-01	+	9.09E-01
cg12555233	15	*MAN2A2*	+	**2.35E-05**	+	**1.63E-03**	+	8.16E-02	+	9.62E-01
cg07558761	16	*SLC7A5*	+	**3.96E-05**	+	**3.37E-02**	+	1.21E-01	+	6.48E-01
cg07021906	16	*SLC7A5*	+	**8.72E-05**	+	**1.10E-02**	+	3.17E-02	+	2.09E-01
cg04583842	16	*BANP*	+	**5.29E-07**	+	**3.97E-05**	+	2.82E-02	+	5.99E-02
cg08857797	17	*VPS25*	+	**4.46E-04**	+	**8.65E-05**	+	8.39E-01	+	7.14E-01
cg22304262	19	*SLC1A5*	-	**4.64E-03**	-	1.79E-01	-	5.54E-01	-	9.50E-01
cg02711608	19	*SLC1A5*	-	**2.52E-02**	-	6.34E-01	-	6.54E-01	+	5.66E-01
cg21766592	19	*SLC1A5*	-	**6.69E-03**	-	4.70E-01	-	8.66E-02	+	9.17E-01
cg06500161	21	*ABCG1*	+	**3.43E-02**	+	**2.98E-02**	+	7.19E-01	-	8.97E-01
cg01820192	21	*C21orf125*	+	**6.41E-04**	+	7.01E-02	+	1.90E-01	+	5.00E-01

Table S3. Associations of the 33 CpG sites with EH in Lifelines DEEP cohort

ProbeID	chr.	Position	Gene	Direction	P value
cg18933331	1	110186418	intergenic	+	0.40982
cg16246545	1	120255941	PHGDH	-	0.59214
cg14476101	1	120255992	PHGDH	-	0.24269
cg19693031	1	145441552	TXNIP	+	0.42444
cg19266329	1	145456128	intergenic	+	0.30504
cg24955196	1	154982621	ZBTB7B	-	0.367741
cg12593793	**1**	**156074135**	**intergenic**	-	**0.00561**
cg01343041	2	24397787	C2orf84	+	0.199573
cg18119407	2	201980504	CFLAR	+	0.07703
cg06690548	4	139162808	SLC7A11	-	0.18537
cg18120259	6	43894639	LOC100132354	-	0.0603
cg21429551	7	30635762	GARS	+	0.96625
cg19390658	7	30636176	GARS	+	0.24793
cg19695041	8	38615330	TACC1	-	0.06736
cg00008629	9	115093661	ROD1	+	0.09652
cg13696706	9	124396830	DAB2IP	+	0.3064
cg11376147	**11**	**57261198**	**SLC43A1**	**-**	**0.001**
cg11468085	11	67435577	ALDH3B2	+	0.26652
cg00574958	11	68607622	CPT1A	-	0.4117
cg00508575	12	90050967	ATP2B1	+	0.38144
cg05248321	14	20898128	KLHL33	-	0.84643
cg00716257	14	75897417	JDP2	+	0.52039
cg02003183	14	103415882	CDC42BPB	+	0.82545
cg26916780	15	64889554	ZNF609	-	0.32459
cg12555233	15	91455366	MAN2A2	+	0.85009
cg07558761	16	87866696	SLC7A5	+	0.69759
cg07021906	16	87866833	SLC7A5	+	0.61261
cg04583842	16	88103117	BANP	+	0.8756
cg08857797	17	40927699	VPS25	-	0.08236
cg22304262	19	47287778	SLC1A5	+	0.14571
cg21766592	**19**	**47288066**	**SLC1A5**	**-**	**0.04095**
cg06500161	**21**	**43656587**	**ABCG1**	**+**	**0.00749**
cg01820192	21	44869762	C21orf125	-	0.7618

Table S4. GSEA results of our meta-analysis (FDR<0.05)

Trait	Biological Process Pathways	FDR
SBP	CELL SUBSTRATE ADHERENS JUNCTION ASSEMBLY	0.029
	CELL SUBSTRATE JUNCTION ASSEMBLY	0.032
	CALCIUM ION IMPORT INTO CYTOSOL	0.039
	ADHERENS JUNCTION ASSEMBLY	0.050
	NEUTRAL AMINO ACID TRANSPORT	0.060
DBP	REGULATION OF GRANULOCYTE DIFFERENTIATION	0.003
	REGULATION OF DENDRITE EXTENSION	0.007
	MEMBRANE BIOGENESIS	0.028
	PROTEIN AUTOPHOSPHORYLATION	0.030
	MEMBRANE ASSEMBLY	0.034
	REGULATION OF PROTEIN AUTOPHOSPHORYLATION	0.042
	NEUTRAL AMINO ACID TRANSPORT	0.074

The pathway previously identified by the CHARGE consortium was highlighted in grey.

Table S5. Heritability of the 34 cross-validated CpG sites in the FTC and NTR cohorts

Probeid	Chr.	Position	Gene	FTC		NTR	
				h^2	95%CI	h^2	95%CI
cg18933331	1	110186418	*intergenic*	0.77	0.69 - 0.83	0.68	0.64 - 0.71
cg16246545	1	120255941	*PHGDH*	0.73	0.63 - 0.80	0.81	0.79 - 0.83
cg14476101	1	120255992	*PHGDH*	0.69	0.57 - 0.77	0.78	0.76 - 0.81
cg19693031	1	145441552	*TXNIP*	0.58	0.46 - 0.68	0.55	0.50 - 0.59
cg19266329	1	145456128	*intergenic*	0.44	0.28 - 0.57	0.35	0.29 - 0.40
cg24955196	1	154982621	*ZBTB7B*	0.53	0.40 - 0.63	0.31	0.24 - 0.37
cg12593793	1	156074135	*intergenic*	0.77	0.69 - 0.84	0.51	0.46 - 0.55
cg01343041	2	24397787	*C2orf84*	0.64	0.52 - 0.73	0.65	0.61 - 0.69
cg18119407	2	201980504	*CFLAR*	0.37	0.22 - 0.51	0.36	0.30 - 0.41
cg06690548	4	139162808	*SLC7A11*	0.31	0.08 - 0.51	0.40	0.35 - 0.46
cg18120259	6	43894639	*LOC100132354*	0.66	0.54 - 0.75	0.60	0.56 - 0.64
cg21429551	7	30635762	*GARS*	0.67	0.55 - 0.75	0.70	0.66 - 0.73
cg19390658	7	30636176	*GARS*	0.37	0.21 - 0.51	0.34	0.28 - 0.40
cg19695041	8	38615330	*TACC1*	0.55	0.41 - 0.66	0.47	0.42 - 0.52
cg00008629	9	115093661	*ROD1*	0.83	0.77 - 0.87	0.72	0.69 - 0.75
cg13696706	9	124396830	*DAB2IP*	0.57	0.44 - 0.67	0.59	0.55 - 0.63
cg11376147	11	57261198	*SLC43A1*	0.50	0.35 - 0.63	0.33	0.27 - 0.39
cg11468085	11	67435577	*ALDH3B2*	0.54	0.41 - 0.65	0.54	0.50 - 0.59
cg00574958	11	68607622	*CPT1A*	0.51	0.38 - 0.62	0.37	0.31 - 0.42
cg00508575	12	90050967	*ATP2B1*	0.54	0.41 - 0.65	0.43	0.37 - 0.48
cg05248321	14	20898128	*KLHL33*	0.71	0.60 - 0.79	0.60	0.56 - 0.64
cg00716257	14	75897417	*JDP2*	0.46	0.30 - 0.60	0.19	0.12 - 0.25
cg02003183	14	103415882	*CDC42BPB*	0.62	0.49 - 0.72	0.62	0.58 - 0.66
cg26916780	15	64889554	*ZNF609*	0.43	0.38 - 0.57	0.37	0.31 - 0.42
cg12555233	15	91455366	*MAN2A2*	0.62	0.51 - 0.71	0.50	0.45 - 0.55
cg07558761	16	87866696	*SLC7A5*	0.61	0.48 - 0.71	0.45	0.40 - 0.50
cg07021906	16	87866833	*SLC7A5*	0.61	0.49 - 0.71	0.61	0.57 - 0.65
cg04583842	16	88103117	*BANP*	0.65	0.53 - 0.74	0.61	0.56 - 0.64
cg08857797	17	40927699	*VPS25*	0.63	0.51 - 0.72	0.33	0.27 - 0.38
cg22304262	19	47287778	*SLC1A5*	0.80	0.73 - 0.85	0.59	0.54 - 0.63
cg02711608	19	47287964	*SLC1A5*	0.72	0.62 - 0.80	0.65	0.61 - 0.68
cg21766592	19	47288066	*SLC1A5*	0.55	0.41 - 0.65	0.60	0.56 - 0.64
cg06500161	21	43656587	*ABCG1*	0.66	0.55 - 0.75	0.56	0.51 - 0.60
cg01820192	21	44869762	*C21orf125*	0.34	0.18 - 0.48	0.39	0.33 - 0.44

Table S6. Separate results of the bivariate SEM analysis for BP and its associated CpG sites in FTC and NTR

Probeid	Gene	Trait	FTC				NTR			
			r_g	r_e	P for dropping pathways		r_g	r_e	P for dropping pathways	
					a_{21}	e_{21}			a_{21}	e_{21}
cg19693031	*TXNIP*	SBP	-0.009(-0.244, 0.253)	-0.203(-0.36, -0.044)	0.943	**0.019**	0.068(-0.088, 0.224)	-0.13(-0.252, -0.003)	0.396	**0.045**
cg13696706	*DAB2IP*	SBP	0.102(-0.149, 0.352)	0.041(-0.132, 0.213)	0.421	0.644	0.042(-0.109, 0.191)	0.068(-0.059, 0.193)	0.587	0.294
cg11468085	*ALDH3B2*	SBP	-0.182(-0.453, 0.067)	0.168(-0.002, 0.327)	0.155	0.053	0.034(-0.111, 0.178)	0.098(-0.031, 0.223)	0.645	0.136
cg05248321	*KLHL33*	SBP	0.13(-0.101, 0.359)	0.029(-0.152, 0.207)	0.266	0.759	0.013(-0.134, 0.157)	0.082(-0.046, 0.206)	0.866	0.208
cg00716257	*JDP2*	SBP	0.023(-0.259, 0.348)	-0.119(-0.288, 0.059)	0.878	0.188	0.042(-0.177, 0.225)	-0.117(-0.238, 0.008)	0.754	0.066
cg04583842	*BANP*	SBP	0.052(-0.189, 0.282)	0.105(-0.073, 0.275)	0.666	0.247	0.086(-0.063, 0.233)	0.01(-0.12, 0.139)	0.255	0.886
cg08857797	*VPS25*	SBP	0.088(-0.147, 0.314)	0.072(-0.102, 0.24)	0.455	0.413	-0.002(-0.186, 0.176)	0.146(0.02, 0.267)	0.983	**0.024**
cg22304262	*SLC1A5*	SBP	-0.133(-0.327, 0.074)	-0.124(-0.294, 0.054)	0.202	0.171	0.017(-0.128, 0.161)	-0.014(-0.144, 0.117)	0.822	0.837
cg06500161	*ABCG1*	SBP	0.109(-0.12, 0.328)	0.088(-0.087, 0.257)	0.345	0.325	-0.07(-0.218, 0.079)	0.061(-0.067, 0.187)	0.355	0.350
cg19693031	*TXNIP*	DBP	0.048(-0.207, 0.329)	-0.148(-0.308, 0.021)	0.722	0.086	-0.041(-0.197, 0.118)	-0.106(-0.231, 0.022)	0.612	0.103
cg11468085	*ALDH3B2*	DBP	-0.167(-0.469, 0.102)	0.147(-0.021, 0.305)	0.229	0.085	0.065(-0.083, 0.211)	0.01(-0.118, 0.139)	0.392	0.875
cg08857797	*VPS25*	DBP	0.227(-0.018, 0.471)	-0.021(-0.191, 0.15)	0.069	0.807	0.055(-0.131, 0.235)	0.106(-0.021, 0.229)	0.557	0.101

r_g: genetic correlation; r_e: unique environmental correlation. A detailed explanation is provided in Supplementary Data Figure S3.

Table S7. Phenotypes previously associated with the 34 CpG sites

CpG sites	Gene	Phenotypes
cg06500161	ABCG1	BMI, Lipid, Type 2 diabetes
cg11468085	ALDH3B2	BMI
cg00508575	ATP2B1	
cg04583842	BANP	BMI, CRP
cg01820192	C21orf125	
cg01343041	C2orf84	
cg02003183	CDC42BPB	CRP
cg18119407	CFLAR	BMI
cg00574958	CPT1A	BMI, lipids
cg13696706	DAB2IP	
cg19390658	GARS	BMI
cg21429551	GARS	BMI
cg00716257	JDP2	Alcohol
cg05248321	KLHL33	CRP
cg18120259	LOC100132354	BMI, A-diol
cg12555233	MAN2A2	CRP
cg14476101	PHGDH	BMI, Lipids, A-diol
cg16246545	PHGDH	BMI, alcohol intake
cg00008629	ROD1	Smoking
cg02711608	SLC1A5	BMI, Lipids, CRP
cg21766592	SLC1A5	BMI
cg22304262	SLC1A5	BMI, A-diol
cg11376147	SLC43A1	BMI
cg06690548	SLC7A11	BMI, Lipids, A-diol
cg07021906	SLC7A5	BMI, CRP
cg07558761	SLC7A5	
cg19695041	TACC1	
cg19693031	TXNIP	BMI, Lipids, Type 2 Diabetes
cg08857797	VPS25	BMI
cg24955196	ZBTB7B	Alcohol
cg26916780	ZNF609	
cg12593793	Intergenic	BMI
cg18933331	Intergenic	diabetes
cg19266329	Intergenic	Alcohol, All-cause mortality

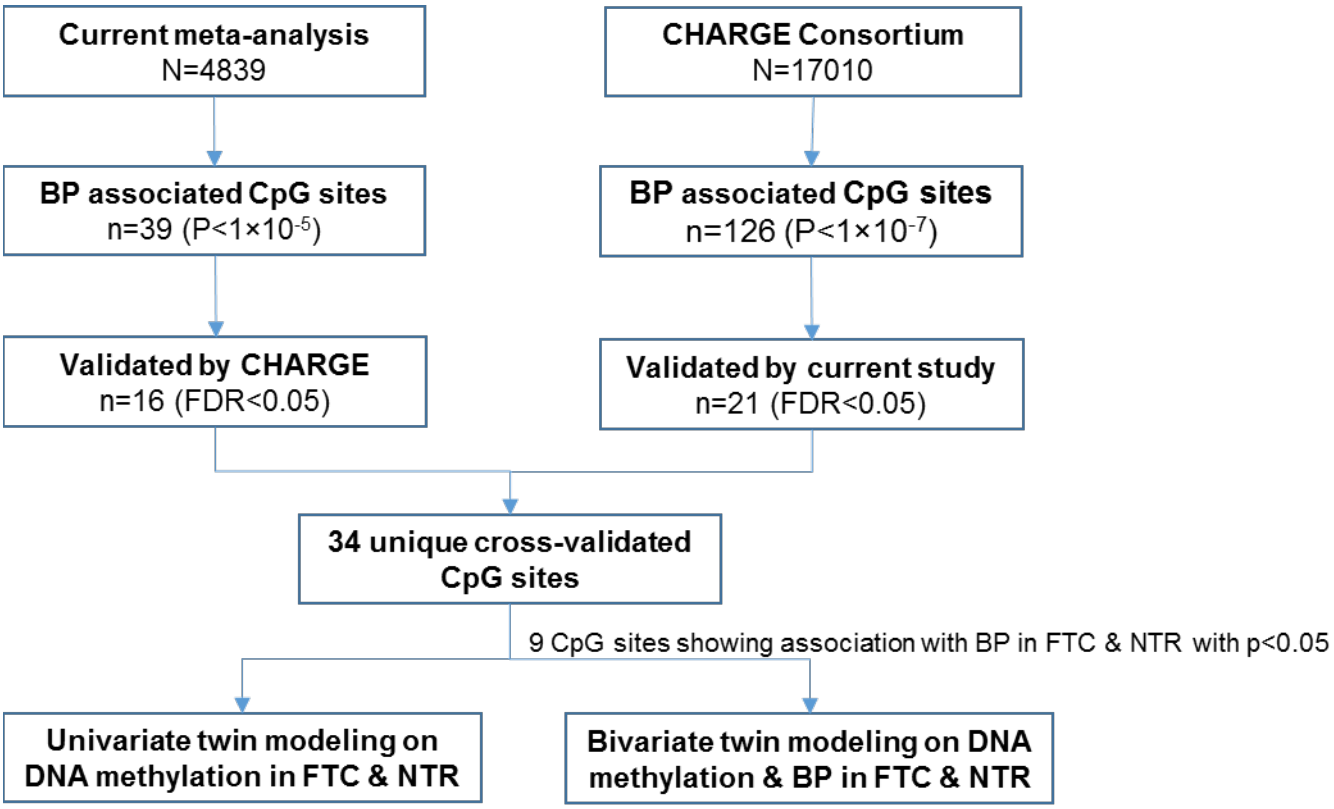
Current meta-analysis
N=4839
CHARGE Consortium
N=17010
BP associated CpG sites
n=39 (P<1×10-5)
BP associated CpG sites
n=126 (P<1×10-7)
Validated by CHARGE
n=16 (FDR<0.05)
Validated by current study
n=21 (FDR<0.05)
34 unique cross-validated CpG sites
9 CpG sites showing association with BP in FTC & NTR with p<0.05
Univariate twin modeling on DNA methylation in FTC & NTR
Bivariate twin modeling on DNA methylation & BP in FTC & NTR

Figure S1.

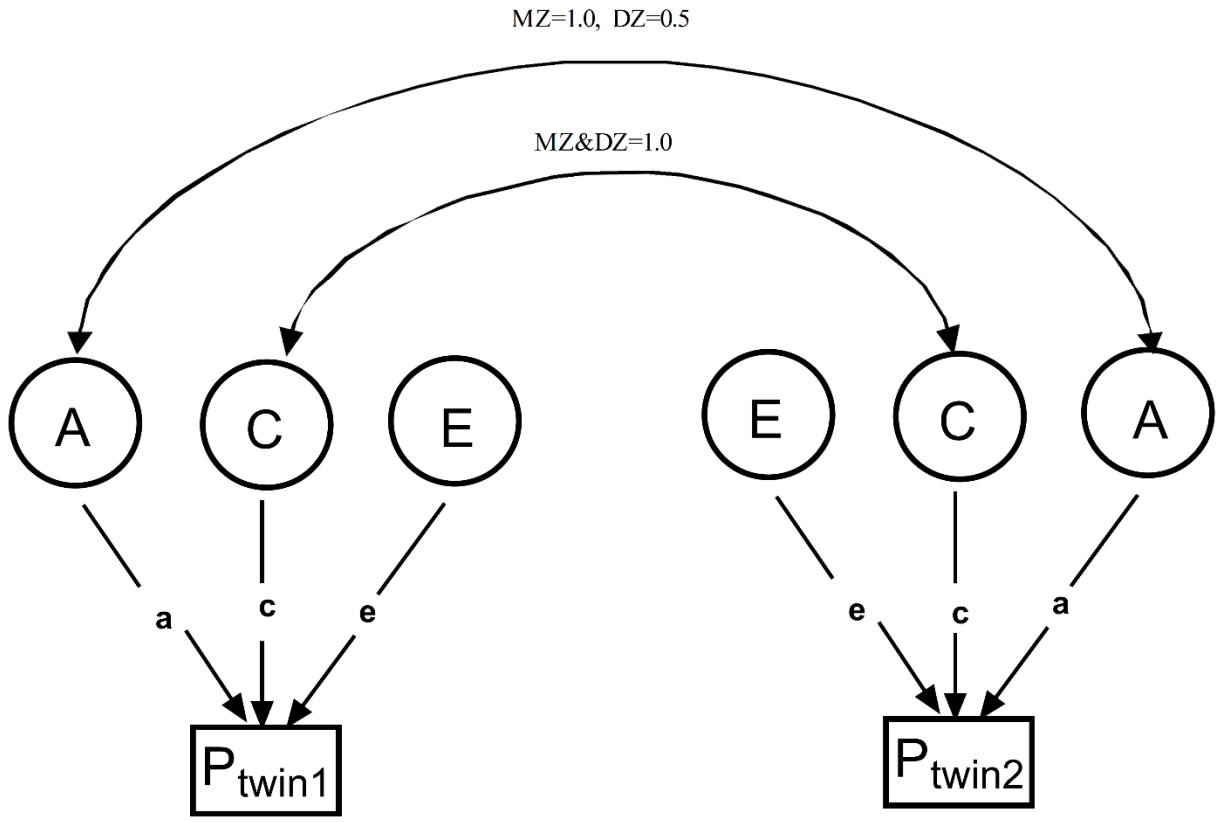

Figure S2

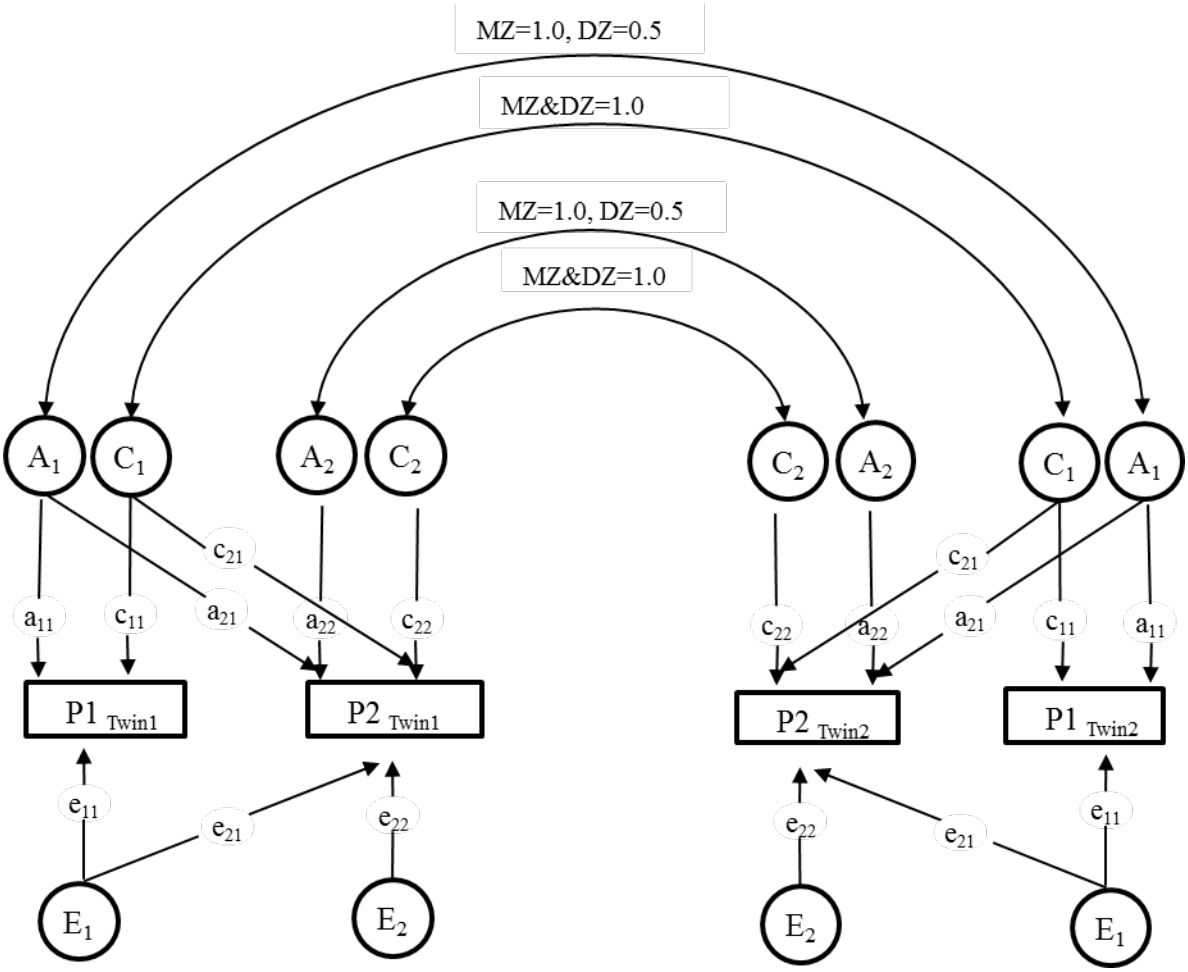
MZ=1.0, DZ=0.5
MZ&DZ=1.0
MZ=1.0, DZ=0.5
MZ&DZ=1.0
A1
C1
A2
C2
C2
A2
C1
A1
c21
c21
a11
c11
a21
a22
c22
c22
a22
a21
c11
a11
P1 Twin1
P2 Twin1
P2 Twin2
P1 Twin2
e11
e21
e22
e22
e21
e11
E1
E2
E2
E1

Figure S3.

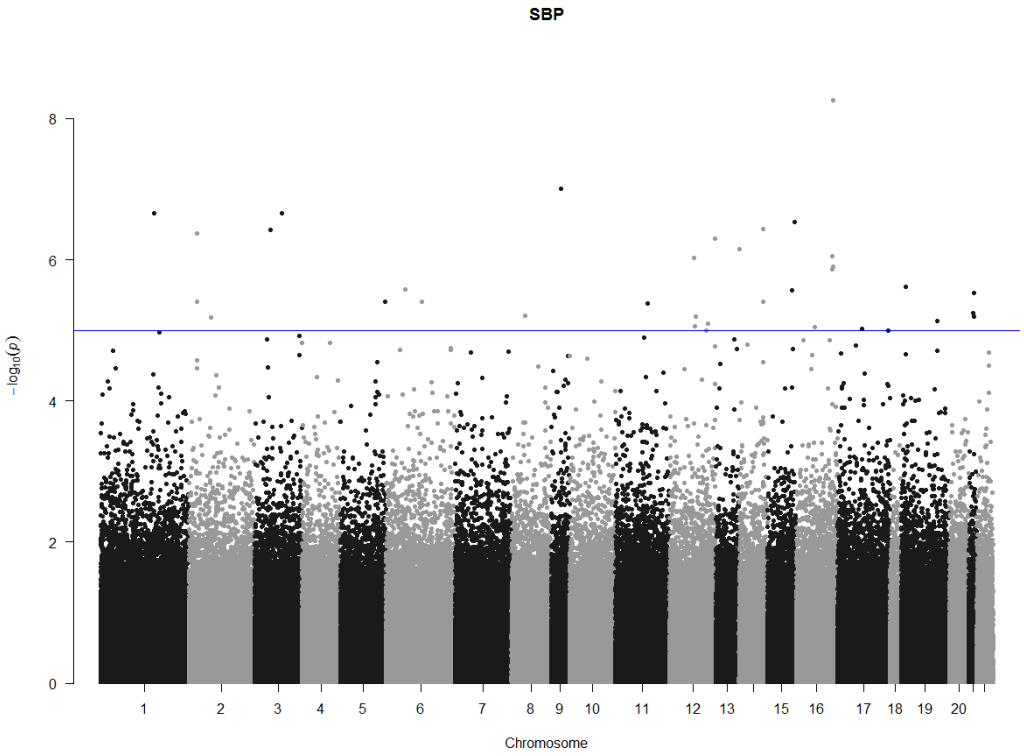

Figure S4a.

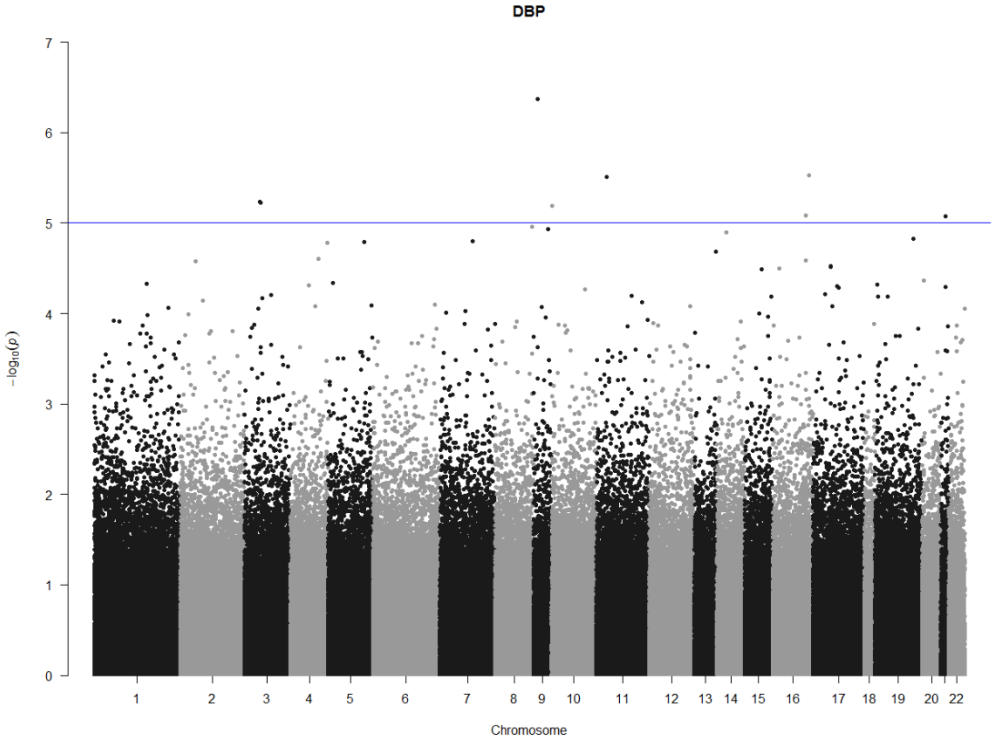

Figure S4b.

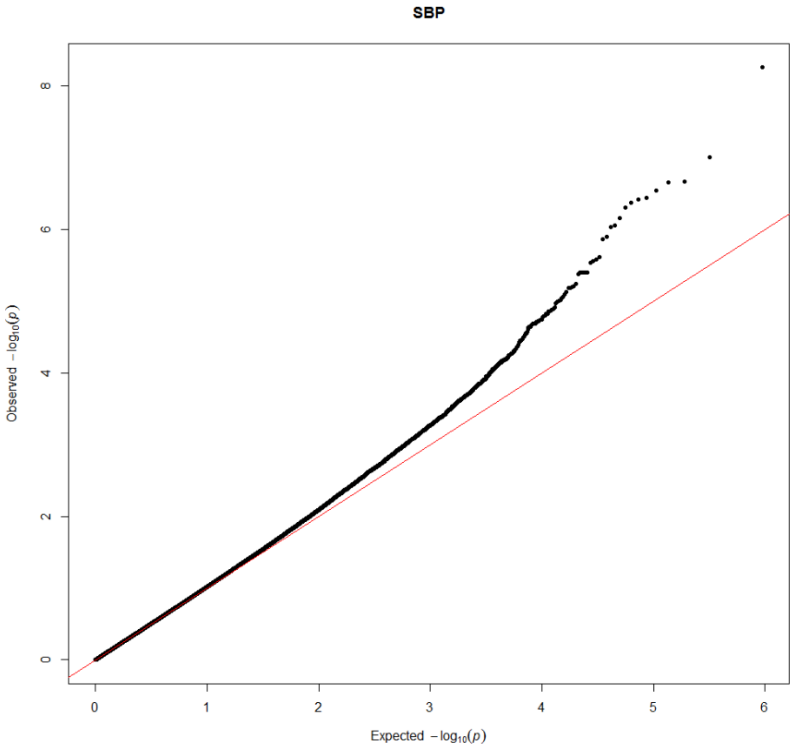

Figure S5a.

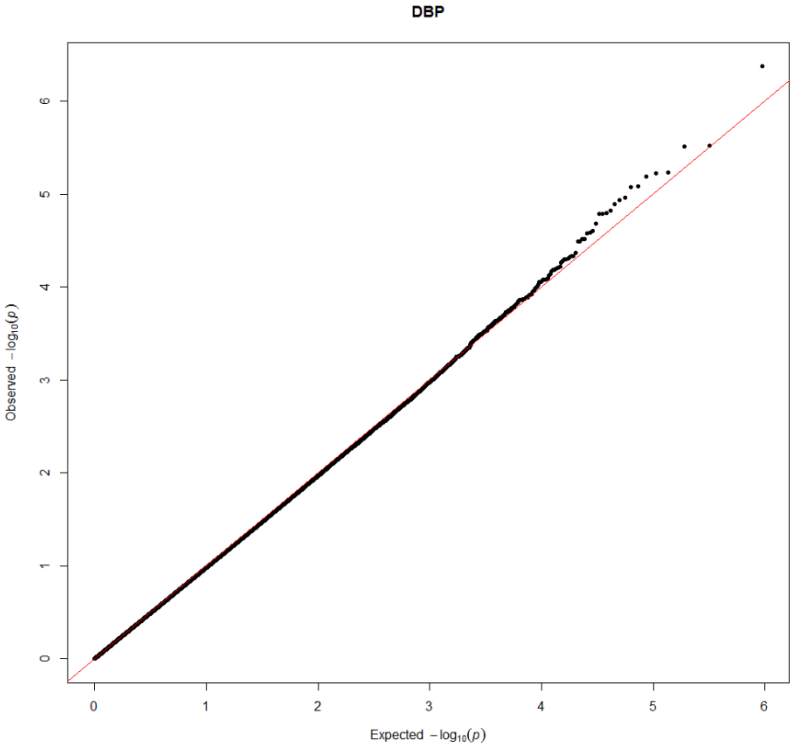

Figure S5b.

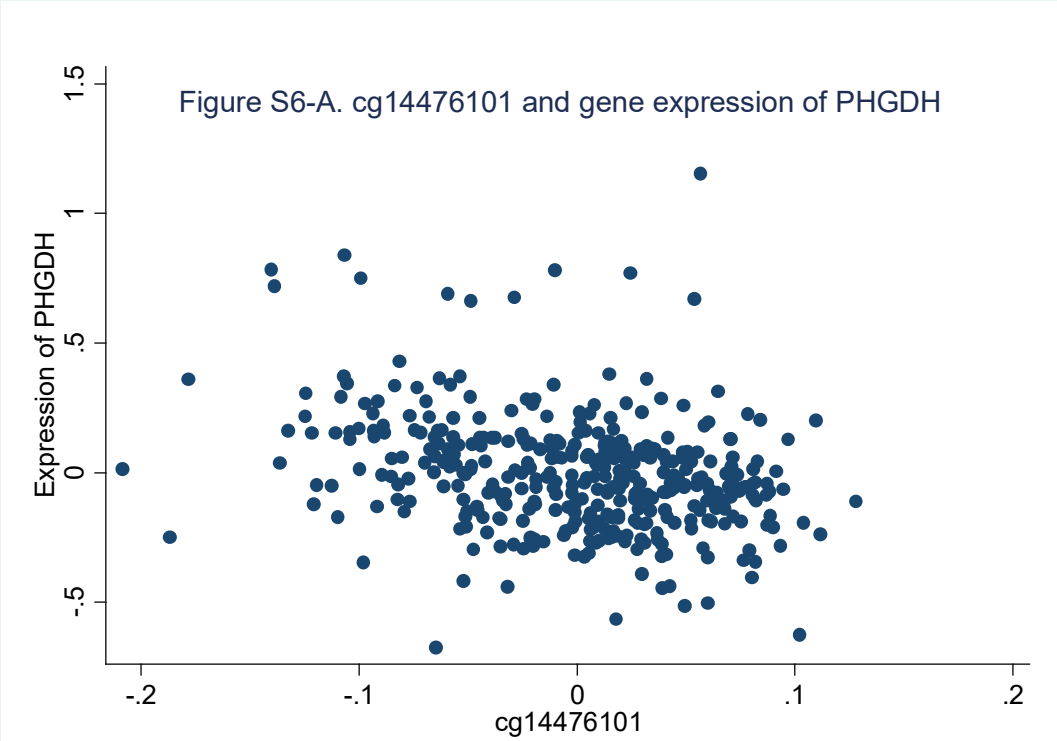
Figure S6-A. cg14476101 and gene expression of PHGDH
Expression of PHGDH
1.5
1
.5
0
-.5
-.2
-.1
0
.1
.2
cg14476101

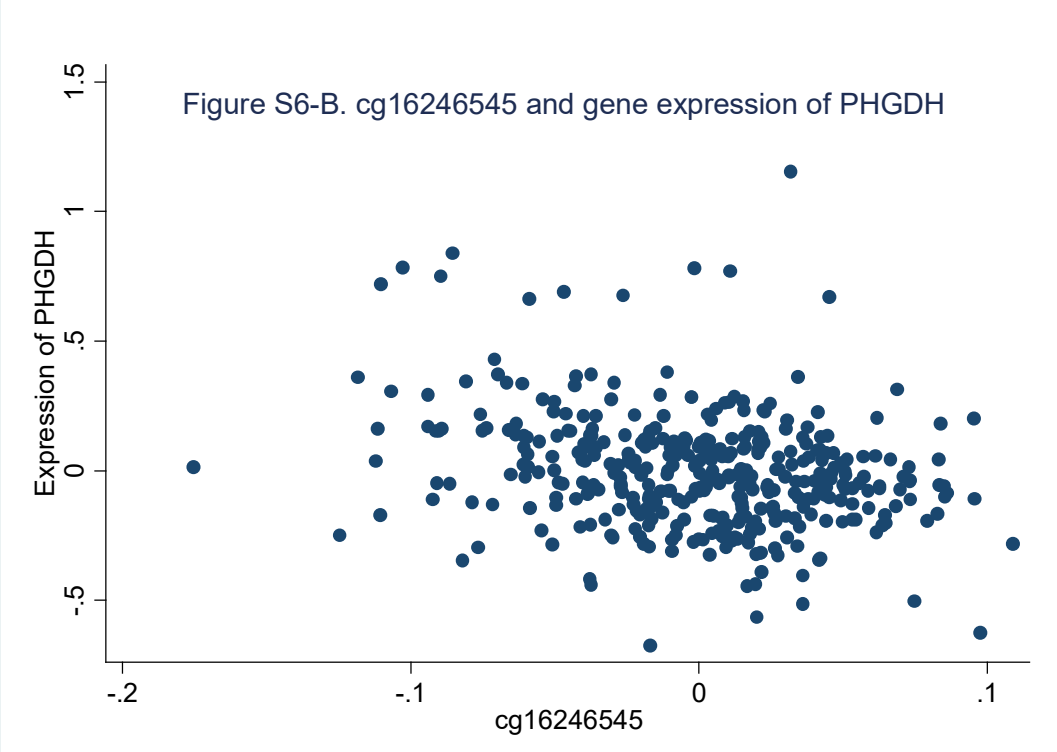
Figure S6-B. cg16246545 and gene expression of PHGDH
Expression of PHGDH
1.5
1
.5
0
-.5
-.2
-.1
0
.1
cg16246545

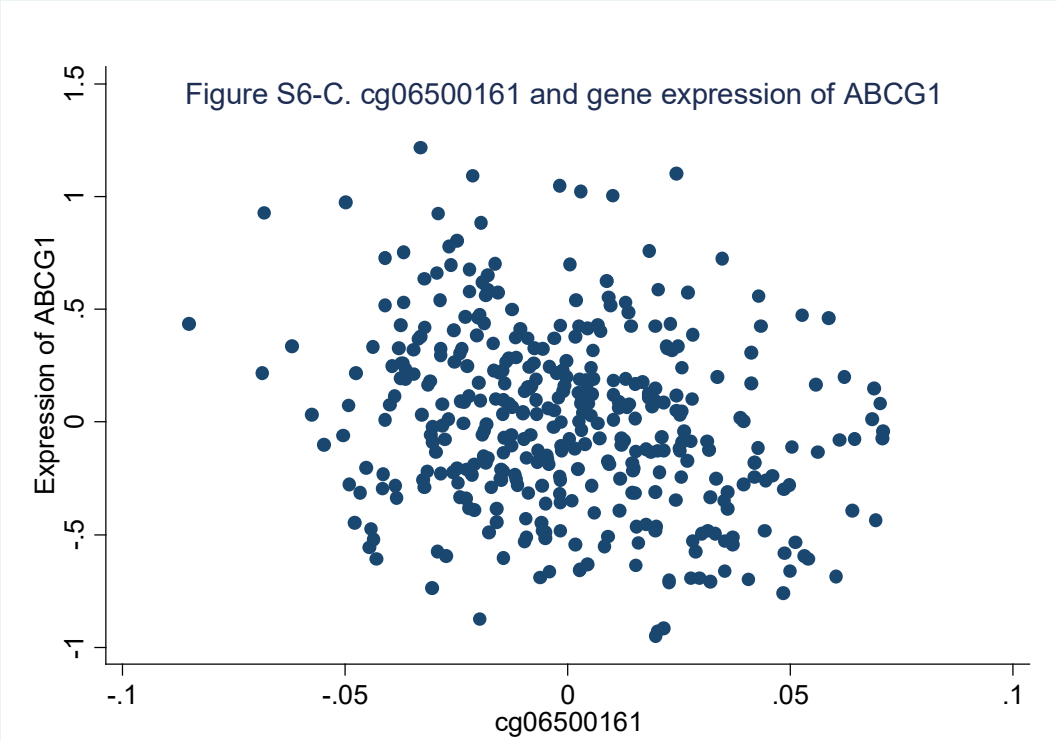
Figure S6-C. cg06500161 and gene expression of ABCG1
Expression of ABCG1
1.5
1
.5
0
-.5
-1
-.1
-.05
0
.05
.1
cg06500161

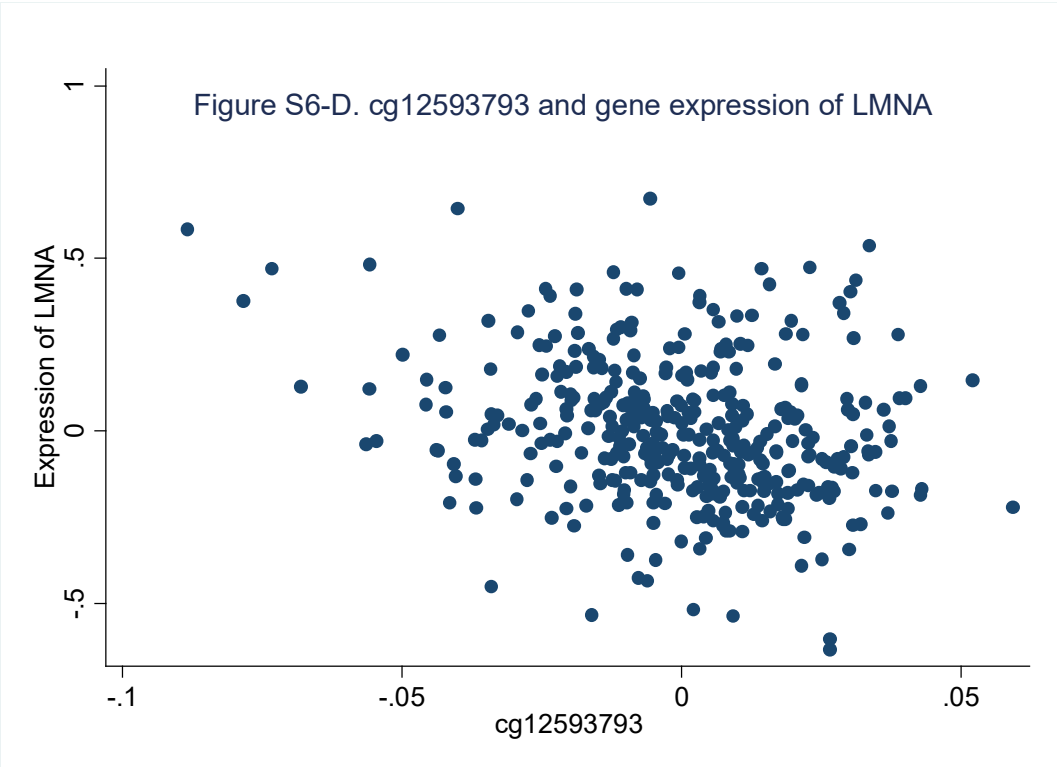
Figure S6-D. cg12593793 and gene expression of LMNA
Expression of LMNA
1
.5
0
-.5
-.1
-.05
0
.05
cg12593793

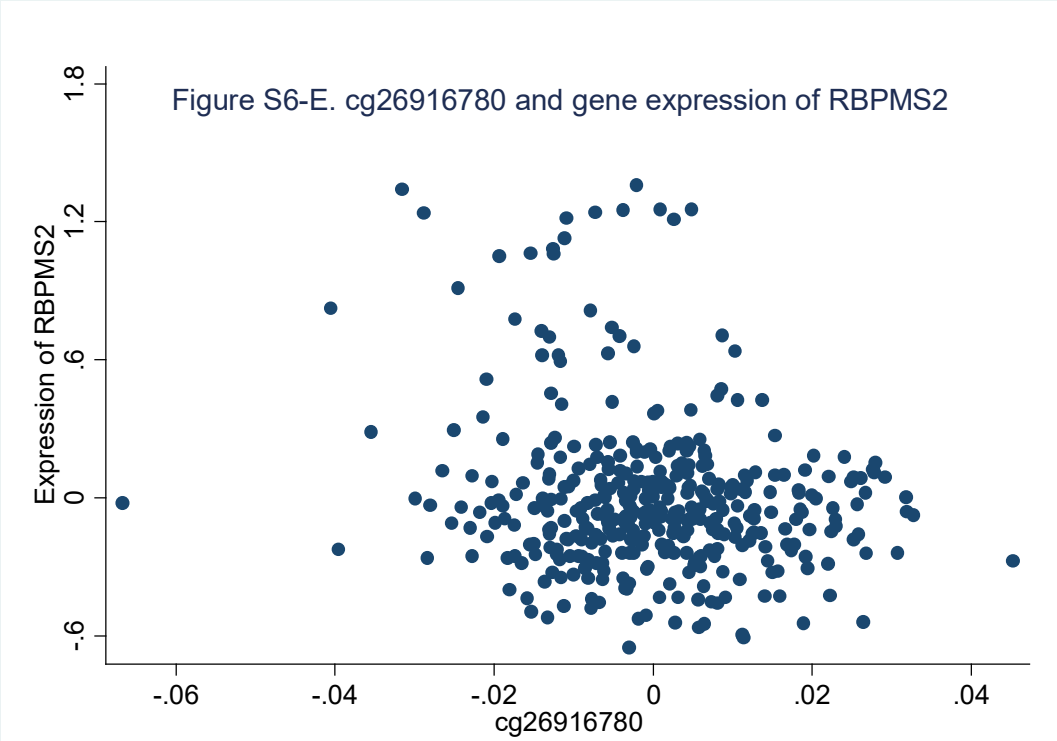
Figure S6-E. cg26916780 and gene expression of RBPMS2
Expression of RBPMS2
1.8
1.2
.6
0
-.6
-.06
-.04
-.02
0
.02
.04
cg26916780

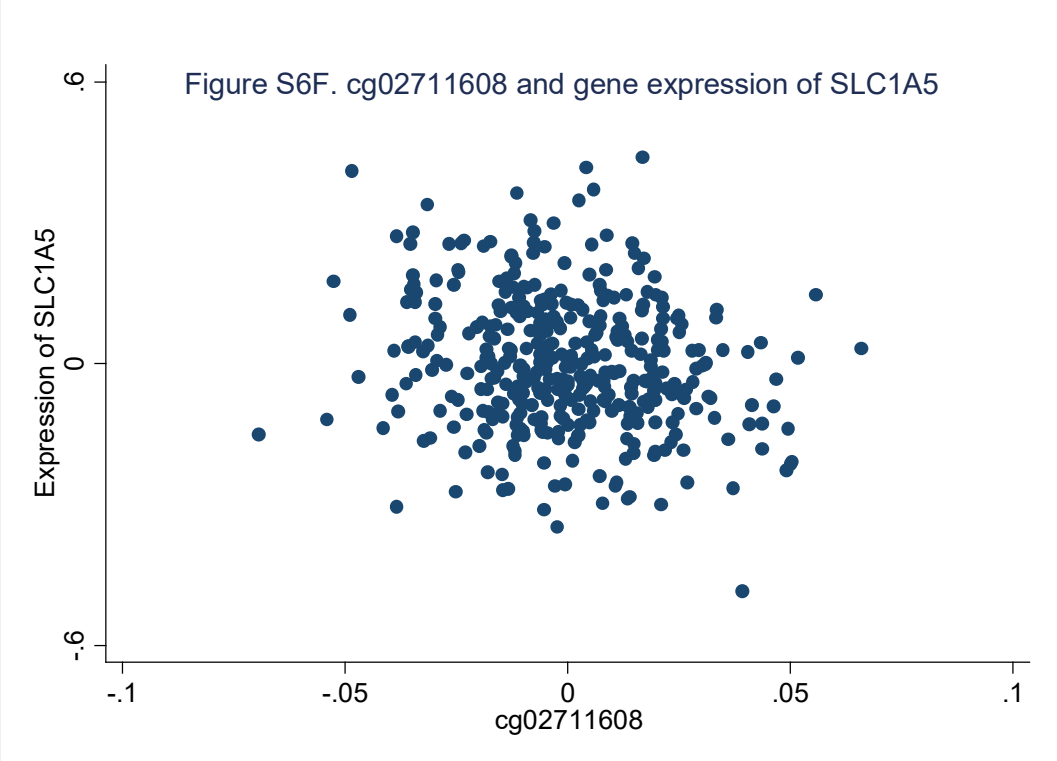
Figure S6F. cg02711608 and gene expression of SLC1A5
Expression of SLC1A5
.6
0
-.6
-.1
-.05
0
.05
.1
cg02711608

REFERENCE

1. Kaprio J. The finnish twin cohort study: An update. *Twin research and human genetics : the official journal of the International Society for Twin Studies*. 2013;16:157-162
2. Tobin MD, Sheehan NA, Scurrah KJ, Burton PR. Adjusting for treatment effects in studies of quantitative traits: Antihypertensive therapy and systolic blood pressure. *Stat Med*. 2005;24:2911-2935
3. Su S, Wang X, Pollock JS, Treiber FA, Xu X, Snieder H, et al. Adverse childhood experiences and blood pressure trajectories from childhood to young adulthood: The georgia stress and heart study. *Circulation*. 2015;131:1674-1681
4. Wang X, Poole JC, Treiber FA, Harshfield GA, Hanevold CD, Snieder H. Ethnic and gender differences in ambulatory blood pressure trajectories: Results from a 15-year longitudinal study in youth and young adults. *Circulation*. 2006;114:2780-2787
5. Wang X, Pan Y, Zhu H, Hao G, Huang Y, Barnes V, et al. An epigenome-wide study of obesity in african american youth and young adults: Novel findings, replication in neutrophils, and relationship with gene expression. *Clinical epigenetics*. 2018;10:3
6. Xu X, Su S, Barnes VA, De Miguel C, Pollock J, Ownby D, et al. A genome-wide methylation study on obesity: Differential variability and differential methylation. *Epigenetics*. 2013;8:522-533
7. Gutin B, Johnson MH, Humphries MC, Hatfield-Laube JL, Kapuku GK, Allison JD, et al. Relationship of visceral adiposity to cardiovascular disease risk factors in black and white teens. *Obesity (Silver Spring)*. 2007;15:1029-1035
8. Taylor HA, Jr., Wilson JG, Jones DW, Sarpong DF, Srinivasan A, Garrison RJ, et al. Toward resolution of cardiovascular health disparities in african americans: Design and methods of the jackson heart study. *Ethnicity & disease*. 2005;15:S6-4-17
9. Pinto-Sietsma SJ, Janssen WM, Hillege HL, Navis G, De Zeeuw D, De Jong PE. Urinary albumin excretion is associated with renal functional abnormalities in a nondiabetic population. *Journal of the American Society of Nephrology : JASN*. 2000;11:1882-1888
10. Pinto-Sietsma SJ, Mulder J, Janssen WM, Hillege HL, de Zeeuw D, de Jong PE. Smoking is related to albuminuria and abnormal renal function in nondiabetic persons. *Annals of internal medicine*. 2000;133:585-591
11. Su S, Miller AH, Snieder H, Bremner JD, Ritchie J, Maisano C, et al. Common genetic contributions to depressive symptoms and inflammatory markers in middle-aged men: The twins heart study. *Psychosomatic medicine*. 2009;71:152-158
12. Klebaner D, Huang Y, Hui Q, Taylor JY, Goldberg J, Vaccarino V, et al. X chromosome-wide analysis identifies DNA methylation sites influenced by cigarette smoking. *Clinical epigenetics*. 2016;8:20
13. Klijs B, Scholtens S, Mandemakers JJ, Snieder H, Stolk RP, Smidt N. Representativeness of the lifelines cohort study. *PloS one*. 2015;10:e0137203
14. Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort profile: Lifelines, a three-generation cohort study and biobank. *International journal of epidemiology*. 2015;44:1172-1180
15. Walaszczyk E, Luijten M, Spijkerman AMW, Bonder MJ, Lutgers HL, Snieder H, et al. DNA methylation markers associated with type 2 diabetes, fasting glucose and hba1c levels: A systematic review and replication in a case-control sample of the lifelines study. *Diabetologia*. 2018;61:354-368
16. Hottenga JJ, Boomsma DI, Kupper N, Posthuma D, Snieder H, Willemsen G, et al. Heritability and stability of resting blood pressure. *Twin research and human genetics : the official journal of the International Society for Twin Studies*. 2005;8:499-508

17. De Geus EJ, Kupper N, Boomsma DI, Snieder H. Bivariate genetic modeling of cardiovascular stress reactivity: Does stress uncover genetic variance? *Psychosomatic medicine*. 2007;69:356-364
18. Mancia G, Parati G. Office compared with ambulatory blood pressure in assessing response to antihypertensive treatment: A meta-analysis. *Journal of hypertension*. 2004;22:435-445
19. Borodulin K, Vartiainen E, Peltonen M, Jousilahti P, Juolevi A, Laatikainen T, et al. Forty-year trends in cardiovascular risk factors in finland. *European journal of public health*. 2015;25:539-546
20. Nelis M, Esko T, Magi R, Zimprich F, Zimprich A, Toncheva D, et al. Genetic structure of europeans: A view from the north-east. *PloS one*. 2009;4:e5472
21. Berenson GS, Wattigney WA, Bao W, Srinivasan SR, Radhakrishnamurthy B. Rationale to study the early natural history of heart disease: The bogalusa heart study. *The American journal of the medical sciences*. 1995;310 Suppl 1:S22-28
22. Eriksson JG, Forsen T, Tuomilehto J, Osmond C, Barker DJ. Early growth and coronary heart disease in later life: Longitudinal study. *Bmj*. 2001;322:949-953
23. Khulan B, Manning JR, Dunbar DR, Seckl JR, Raikkonen K, Eriksson JG, et al. Epigenomic profiling of men exposed to early-life stress reveals DNA methylation differences in association with current mental state. *Translational psychiatry*. 2014;4:e448
24. Kananen L, Marttila S, Nevalainen T, Jylhava J, Mononen N, Kahonen M, et al. Aging-associated DNA methylation changes in middle-aged individuals: The young finns study. *BMC genomics*. 2016;17:103
25. Oikonen M, Nuotio J, Magnussen CG, Viikari JS, Taittonen L, Laitinen T, et al. Repeated blood pressure measurements in childhood in prediction of hypertension in adulthood. *Hypertension*. 2016;67:41-47
26. Tigchelaar EF, Zhernakova A, Dekens JA, Hermes G, Baranska A, Mujagic Z, et al. Cohort profile: Lifelines deep, a prospective, general population cohort study in the northern netherlands: Study design and baseline characteristics. *BMJ open*. 2015;5:e006772
27. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: A flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. *Bioinformatics*. 2014;30:1363-1369
28. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the illumina humanmethylation450 beadchip improves data quality and performance in epigenome-wide association studies. *Genome biology*. 2015;16:37
29. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86
30. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing illumina 450k methylation array data. *BMC genomics*. 2013;14:293
31. van Iterson M, Tobi EW, Slieker RC, den Hollander W, Luijk R, Slagboom PE, et al. Methylaid: Visual and interactive quality control of large illumina 450k datasets. *Bioinformatics*. 2014;30:3435-3437
32. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic cpgs in the illumina infinium humanmethylation450 microarray. *Epigenetics*. 2013;8:203-209
33. Genome of the Netherlands C. Whole-genome sequence variation, population structure and demographic history of the dutch population. *Nature genetics*. 2014;46:818-825
34. Bonder MJ, Kasela S, Kals M, Tamm R, Lokk K, Barragan I, et al. Genetic and epigenetic regulation of gene expression in fetal and adult human livers. *BMC genomics*. 2014;15:860

35. Touleimat N, Tost J. Complete pipeline for infinium((r)) human methylation 450k beadchip data processing using subset quantile normalization for accurate DNA methylation estimation. *Epigenomics*. 2012;4:325-341
36. Zhuang J, Widschwendter M, Teschendorff AE. A comparison of feature selection and classification methods in DNA methylation studies using the illumina infinium platform. *BMC Bioinformatics*. 2012;13:59
37. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical bayes methods. *Biostatistics*. 2007;8:118-127